**Curtin School of Allied Health**

# Exploring the Application of Wearable Movement Sensors in People with Knee Osteoarthritis

**Jay-Shian Tan**

**ORCID ID: 0000-0002-9728-3128**

**This thesis is presented for the Degree of**
**Doctor of Philosophy**
**of**
**Curtin University**

**October 2022**

# Author's Declaration

I declare that this thesis is my own account of my research and contains work which has not previously been submitted for a degree at any tertiary education institution. To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

Human Ethics: The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2018. The proposed research study received human research ethics approval from the Curtin Human Research Ethics Committee (approval number: HRE2017-0738).

**Jay-Shian Tan**

28 October 2022

# Statement of Contributors

The candidate, Jay-Shian Tan, was responsible for all aspects of the research presented within this thesis, including study design, ethics approval, data analysis, interpretation, reporting of results and process of submitting manuscripts. The following supervisory team and co-authors also contributed to the research design, and some aspects of analysis, interpretation, and writing/editing.

| | |
|---|---|
| **Jay-Shian Tan** | October 2022 |

I, as co-author, endorse that this level of contribution indicated by the candidate.

| | |
|---|---|
| Associate Professor Peter Kent | October 2022 |
| Professor Anne Smith | October 2022 |
| Associate Professor Amity Campbell | October 2022 |
| Professor Peter O'Sullivan | October 2022 |
| Dr J.P. Caneiro | October 2022 |
| Tara Binnie | October 2022 |
| Paul Davey | October 2022 |
| Dr Sawitchaya Tippaya | October 2022 |
| Dr Kathryn Napier | October 2022 |
| Dr Behrouz Beheshti | October 2022 |
| Edward Tikoft | October 2022 |

# Abstract

Knee osteoarthritis accounts for 83% of the total osteoarthritis burden and is a growing musculoskeletal health problem. The condition significantly impacts health care systems, the workforce, and most importantly, individuals living with this condition. A contemporary understanding of knee osteoarthritis considers this condition in the context of the whole person. Symptoms of knee osteoarthritis can include joint pain, stiffness, sleep disturbance, lower limb weakness, altered movement patterns and emotional distress. The condition is associated with physical impairment and activity limitation that affect a person's ability to participate in society. The most common activity limitations include walking, negotiating stairs, and transitioning to and from a chair. Some people with knee osteoarthritis avoid or reduce how often they engage in such an activity because it is painful or difficult to perform.

Across a range of activities, people with knee osteoarthritis demonstrate altered movement patterns affecting kinematics, kinetics, and muscle activity. During activities such as walking, negotiating stairs, and transitioning to and from a chair, there are biomechanical parameters (movement parameters) that differentiate people with and without knee osteoarthritis. Some of these movement parameters include reduced knee flexion angles, increased knee adduction and flexion moments, as well as quadriceps-hamstrings co-contraction. Various movement parameters have been associated with increased risk of structural progression, activity limitation or pain as well as cognitive factors such as fear and reduced confidence. However, it is unclear if changes in movement patterns after an intervention are related to changes in clinical outcomes such as pain and activity limitation, and clinicians do not currently have a practical means of monitoring movement patterns outside of a clinical environment.

The broad aims of the research in this doctoral thesis are to: (1) explore the relationship between a change in movement patterns and change in clinical outcomes following exercise interventions, and (2) investigate how wearable sensor technology could be used to monitor activity avoidance and altered movement patterns in people with knee osteoarthritis.

## Study 1 – Systematic Review

**Background:** Exercise is recommended in many clinical guidelines as a core intervention for people with knee osteoarthritis. Some types of exercise interventions, such as neuromuscular exercise and gait retraining, directly target changes in movement patterns, while other exercise interventions, including resistance training or general walking programmes, do not. All these various types of exercise have been shown to improve activity limitation and pain in randomised controlled trials. Exercise has the potential to modify movement patterns directly through changes in technique or indirectly through changes in muscle activity or reduced symptoms. While exercise has the potential to change movement patterns and improve clinical outcomes such as activity limitation or pain, it is unclear the extent to which there is a relationship between those two types of changes.

**Aim:** To determine if there is a relationship between changes in movement patterns and activity limitation or pain after exercise interventions in people with knee osteoarthritis.

**Methods:** Prospectively registered (PROSPERO CRD42020160164) systematic review using the PRISMA statement for reporting standards. *Search*: Four databases (MEDLINE, Embase, CINAHL and AMED) were searched up to 22 January 2021. *Inclusion*: Cohort studies and randomised controlled trials investigating exercise interventions for people with knee osteoarthritis that assessed changes in knee joint movement parameters and changes in the clinical outcomes of activity limitation or pain. *Selection*: Two reviewers independently examined titles, abstracts and full texts using Covidence. *Outcomes*: Changes in kinematic, kinetic and muscle activity parameters during functional activities and activity limitation or pain. *Quality*: Methodological quality was assessed using an adapted version of the Joanna Briggs Institute Risk of Bias Critical Appraisal Checklist and overall quality of evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) Tool. *Data analysis*: A descriptive synthesis of studies that reported group mean change for both movement parameters and clinical outcomes, or a correlation between those outcomes.

**Results:** Data from 22 mostly low-quality studies were analysed involving 936 participants. Eight knee joint moment, 15 kinematic and four muscle activity walking related movement parameters were investigated. No other activities had been investigated. Group mean change was reported in 20 studies for movement parameters and activity limitation or pain, and two studies reported correlations. There was a group-level change relationship (co-occurrence of change in both movement parameter and clinical outcome) 24.5% of the time a movement parameter was investigated, despite an improvement in clinical outcomes occurring 90% of the time. There was an individual-level change relationship in one of eight correlations tested across two studies.

**Discussion:** There was an infrequent relationship between changes in walking-related movement parameters and changes in activity limitation or pain across studies. The exact nature of a change relationship is unclear because of the limited number of high-quality studies, methodological issues, limited biomechanical assessment of activities other than walking, and heterogeneous participant characteristics (e.g. structure, symptoms, behaviour, and cognitions). Collectively the findings suggest that: changes in walking-related movement parameters may be unrelated to improvement in clinical outcomes, or clinical phenotypes may exist where movement parameters are more clinically relevant and responsive to exercise interventions, in which case movement parameters and clinical outcomes would need to be more individualised to accommodate that heterogeneity, or exercise interventions may need to be more targeted.

**This study has been published:**

**Tan, J.-S.**, Tikoft, E., O'Sullivan, P., Smith, A., Campbell, A., Caneiro, J. P., & Kent, P. (2021). The Relationship between Changes in Movement and Activity Limitation or Pain in People with Knee Osteoarthritis: A Systematic Review. *Journal of Orthopaedic & Sports Physical Therapy, 51*(10), 492-502. https://doi.org/10.2519/jospt.2021.10418

## Study 2a – Human Activity Recognition

**Background:** People with knee osteoarthritis often avoid activities and demonstrate altered movement patterns that result in increased mechanical loading. Monitoring activity avoidance and altered movement patterns might be of clinical relevance in a proportion of the population with knee osteoarthritis. However, there are no systems that are accessible for clinicians to monitor activity avoidance or altered movement patterns of their patients in free-living environments (e.g. at home or work). Wearable sensor systems, such as inertial measurement units (IMUs), are a technology that could assist with monitoring patients in free-living environments. However, IMU data streams are long and unlabelled, and therefore do not provide contextual information about what activity a person is doing. Knowing what a person is doing is important for clinicians to make clinically informed decisions about activity avoidance and for meaningful biomechanical analysis. Machine learning, a form of artificial intelligence, is one data handling approach that can be used to process IMU data.

**Aim:** To develop a human activity recognition system to classify clinically important activities (walking, negotiating stairs and transitioning to and from a chair), and phases of activities using raw IMU training data from people with knee osteoarthritis.

**Methods:** IMU data was collected from 18 participants with knee osteoarthritis performing clinically important activities. Convolutional neural network (machine learning) models were trained to predict three levels of classification – activity, direction, and phase.

**Results:** The model accuracy was 85% at the first level of classification (*activity* – walk, stair, chair), 89% to 97% at the second (*direction of movement* – stand to sit/sit-to-stand, ascend/descend stairs) and 60% to 67% at the third level (*swing or stance phase*) for walking and ascending/descending stairs).

**Discussion:** This study was the first to develop a machine learning system for human activity recognition using IMU data collected from people with knee osteoarthritis. Clinically, data from the first and second levels of classification have

the potential to be used to monitor activity avoidance in people with knee osteoarthritis. Labelled data from the second and third level of classification is potentially suitable for use as part of a data handling pipeline for subsequent biomechanical analysis.

**This study has been published:**

**Tan, J.-S.**, Beheshti, B. K., Binnie, T., Davey, P., Caneiro, J. P., Kent, P., . . . Campbell, A. (2021). Human Activity Recognition for People with Knee Osteoarthritis—a Proof-of-Concept. *Sensors, 21*(10). https://doi.org/10.3390/s21103381

## Study 2b – Sagittal Plane Angular Prediction

**Background:** When walking, negotiating stairs and transitioning to and from a chair, people with knee osteoarthritis often use less knee flexion. IMU systems typically integrate data from an accelerometer, gyroscope, and magnetometer using fusion algorithms to output meaningful information like knee flexion angles. However, in free-living environments, electromagnetic interference can affect magnetometer data and impact the validity of kinematic outputs. As an alternative to the use of fusion algorithms and magnetometer data, machine learning has been used to handle raw data from IMU's accelerometers and gyroscopes to predict knee joint kinematics in young, healthy people for the activity of walking. No studies have yet trained a machine learning model on IMU data collected from people with knee osteoarthritis to predict sagittal plane kinematics for multiple clinically important activities using data collected from people with knee osteoarthritis.

**Aims:** To (a) develop machine learning models for the prediction of sagittal plane angular kinematics for multiple clinically important activities using IMU data collected from people with knee osteoarthritis, and (b) explore the model performance of a single-leg model (two IMUs on a single leg), compared to a double-leg model (four IMUs across two legs).

**Methods:** Simultaneous IMU and Vicon data were collected from 17 participants with knee osteoarthritis performing clinically important activities. Bidirectional long-short term memory (machine learning) models were trained to predict sagittal plane angular kinematics for phases of walking, negotiating stairs, and transitioning to and from a chair.

**Results:** The single-leg model was more accurate than the double-leg model for walking and negotiating stairs, while the double-leg model was more accurate for transitioning to and from a chair. The prediction error for the single-leg model ranged from RMSE = 7° to 11° and Pearson's r = 0.89 to 0.99. The prediction error for the double-leg model ranged from RMSE 7° to 13° and Pearson's r = 0.74 to 0.99.

**Discussion:** This study was the first to develop kinematic prediction models for multiple clinically relevant activities for people with knee osteoarthritis. The prediction model was designed to be incorporated into a larger IMU data handling pipeline. Kinematic prediction would be preceded by the labelling of activities and activity-phases of multiple clinically important activities using a human activity recognition algorithm. This study also demonstrated that performance of single-leg and double-leg models may depend on the type of activity being performed. The double-leg model outperformed the single-leg model for symmetrical activities (e.g. sit-to-stand), whereas the single-leg model outperformed the double-leg model for asymmetrical activities (e.g. walking and negotiating stairs). The number of required IMUs used for training each model has implications for patient and clinician burden.

**This study has been published:**

**Tan, J.-S.**, Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., . . . Campbell, A. (2022). Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models. *Sensors, 22*(2). https://doi.org/10.3390/s22020446

## Study 2c – Moment and Force Prediction

**Background:** An increase in frontal and sagittal plane loading during the stance phase of walking is associated with structural progression of medial knee osteoarthritis and increased pain. Knee joint moment and forces have the potential to change in some people with knee osteoarthritis in response to biomechanical interventions such as exercise. Yet, clinicians do not currently have the ability to objectively assess knee joint loading in free-living environments.

Movement patterns across the population with knee osteoarthritis are diverse. This heterogeneity may undermine the model performance of machine learning based biomechanical predictions for population-based models tested on an individual. Machine learning models are commonly tested on participants whose data are not included in training the model. However, individualising biomechanical prediction models may help to accommodate heterogeneity in a clinical environment. One method of individualising prediction models involves adding some of the test participants data to the training dataset (that includes all other study participants), resulting in a more individualised model. It is also unclear if training a machine learning model on single-leg or double-leg data for the prediction of knee moments and forces might affect model performance.

**Aims:** To (a) develop machine learning models for the prediction of knee joint moments and forces for the stance phase of walking from IMU data collected from people with knee osteoarthritis, (b) explore model performance at the level of individual participants when training data from that participant is added to the training phase of a model trained on all other participants, and (c) explore performance of a single-leg model compared to a double-leg model.

**Methods:** IMU and Vicon data were collected from 17 participants with knee osteoarthritis performing clinically important activities. Bidirectional long-short term memory (machine learning) models were trained to predict knee adduction and flexion moments as well as compression and medial force for the stance phase of walking.

**Results:** Individualised models outperformed non-individualised models by normalised RMSE values of 9% to 36%. Double-leg models outperformed single-leg models by normalised RMSE 1% to 23%. The strongest model performance (individualised double-leg model) ranged from normalised RMSE 16% to 23%, Pearson's r = 0.80 to 0.86.

**Discussion:** Individualising prediction models improved machine learning model prediction performance for knee moments and forces in a population with knee osteoarthritis. Double-leg models outperformed single-leg models, suggesting that for optimal results, kinematic and kinetic prediction models may differ in IMU placement requirements across lower limbs. The kinetic prediction models were designed to be used secondary to human activity recognition and may assist clinicians to monitor patients who demonstrate biomechanical loading risk factors related to structural progression of knee osteoarthritis. If individualised prediction models were to be implemented in clinical practice, there are some options about of how this might occur without access to an appropriate motion analysis reference standard.

**This study is under review.**

## Conclusion

This doctoral thesis adds knowledge to the current understanding of the relationship between movement patterns and clinical outcomes in people with knee osteoarthritis, and how movement data from IMU technology can be processed with machine learning algorithms for clinically relevant applications. At a group level, a change in movement patterns and clinical outcomes after exercise interventions were infrequently related. However, there were several methodological limitations in the existing literature that precluded a deeper understanding of the relationship between how people move and their clinical outcomes. There remains the possibility that for a biomechanical phenotype a change in movement pattern is clinically important, but this has not yet been tested thoroughly.

Human activity recognition and biomechanical prediction models were developed as a proof-of-concept for the development of an automated data handling pipeline for IMU data collected from people with knee osteoarthritis. These models were developed to predict clinically important activities and movement parameters. As the foundation of a data handling pipeline, the models could be integrated into a system that provides information about activity avoidance or altered movement patterns in free-living environments. The research into human activity recognition and biomechanical prediction for clinical populations is in its early stages and there are substantial clinical and machine learning considerations that require attention prior to clinical implementation. Sensor based technology together with machine learning has significant potential to assist with patient monitoring to assess physical function, assist telehealth, provide automated biofeedback, and enhance biomechanical risk prediction. Such systems could facilitate clinical research to help establish the relationship between changes in biomechanics and clinical outcomes at an individual person-level in response to intervention.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations and Terms

| Abbreviation | Term |
|---|---|
| ANN | Artificial Neural Network |
| BiLSTM | Bidirectional Long-Short Term Memory |
| BMI | Body Mass Index |
| CI | Confidence Interval |
| CNN | Convolutional Neural Network |
| HAR | Human Activity Recognition |
| IMU | Inertial Measurement Unit |
| KAM | Knee Adduction Moment |
| KOOS | Knee Injury and Osteoarthritis Outcome Score |
| LOSOCV | Leave-One-Subject-Out Cross-Validation |
| LOOCV | Leave-One-Out Cross-Validation |
| LSTM | Long-Short Term Memory |
| MAE | Mean Absolute Error |
| nRMSE | Normalised Root Mean Square Error |
| OARSI | Osteoarthritis Research Society International |
| r | Pearson's correlation coefficient |
| $R^2$ | Coefficient of determination |
| RMSE | Root Mean Square Error |
| SD | Standard Deviation |
| SMD | Standardised Mean Difference |
| SRM | Standardised Response Mean |
| WHR | Wait-to-Hip Ratio |
| WOMAC | Western Ontario and McMaster Universities Osteoarthritis Index |
| VAS | Visual Analogue Scale |

# Thesis Publications

Chapter 3

**Tan, J.-S.**, Tikoft, E., O'Sullivan, P., Smith, A., Campbell, A., Caneiro, J. P., & Kent, P. (2021). The Relationship Between Changes in Movement and Activity Limitation or Pain in People with Knee Osteoarthritis: A Systematic Review. *Journal of Orthopaedic & Sports Physical Therapy, 51*(10), 492-502. https://doi.org/10.2519/jospt.2021.10418

Chapter 4

**Tan, J.-S.**, Beheshti, B. K., Binnie, T., Davey, P., Caneiro, J. P., Kent, P., Smith, A., O'Sullivan, P., & Campbell, A. (2021). Human Activity Recognition for People with Knee Osteoarthritis—A Proof-of-Concept. *Sensors, 21*(10). https://doi.org/10.3390/s21103381

Chapter 5

**Tan, J.-S.**, Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., Kent, P., Smith, A., O'Sullivan, P., & Campbell, A. (2022). Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models. *Sensors, 22*(2). https://doi.org/10.3390/s22020446

# Submitted Manuscripts

Chapter 6

**Tan, J.-S.**, Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., Smith, A., O'Sullivan, P., Campbell, A., & Kent, P. (2022). Deep Learning for Predicting Moments and Forces from Wearable Sensors in People with Knee Osteoarthritis

# Co-authored Publications

Binnie, T., Smith, A., Kent, P., Ng, L., O'Sullivan, P., **Tan, J.-S.**, Davey, P., & Campbell, A. (2021). Concurrent Validation of Inertial Sensors for Measurement of Knee Kinematics in Individuals with Knee Osteoarthritis: A Technical Report. *Health and Technology*, 1-10. https://doi.org/https://doi.org/10.1007/s12553-021-00616-9

Wernli, K., **Tan, J.-S.**, O'Sullivan, P., Smith, A., Campbell, A., & Kent, P. (2021). The Relationship Between Changes in Movement and Changes in Low Back Pain: A Systematic Review of Single-Case Designs. *JOSPT Cases, 1*(4), 199-219. https://doi.org/10.2519/josptcases.2021.10231

Wernli, K., **Tan, J.-S.**, O'Sullivan, P., Smith, A., Campbell, A., & Kent, P. (2020). Does Movement Change When Low Back Pain Changes? A Systematic Review. *Journal of Orthopaedic & Sports Physical Therapy, 50*(12), 664-670. https://doi.org/10.2519/jospt.2020.9635

# Presentations

*Curtin University, Physiotherapy and Exercise Science Emerging Research Conference 2020*

The relationship between changes in movement and activity limitation or pain in people with knee osteoarthritis: A systematic review of exercise interventions

*OPUS Centre for Research Excellence in Total Joint Replacement Forum 2021*

Developing a machine learning system for activity recognition in people with knee osteoarthritis

*OPUS Centre for Research Excellence in Total Joint Replacement Research Day 2022*

Machine learning and wearable movement sensors in people with knee osteoarthritis

# Prizes and Awards

*Physiotherapy Research Foundation 'Best Presentation'*

Curtin University, Physiotherapy and Exercise Science Emerging Research Conference 2020

# Acknowledgements

Thank you to the amazing team of supportive supervisors, co-authors, colleagues, friends, and family members who have helped me in my PhD journey. I have learned so much from you all.

I could not have asked for a more perfect primary supervisor. How fortunate I was that Peter Kent joined the team when I was considering commencing my journey. Always open, honest, patient, accessible, kind, and caring, your nurturing approach to supervision helped me grow my critical thinking and communication skills in ways I could not have imagined. I learned so much from your mentoring and will continue to apply that that knowledge in my own mentoring and supervision into the future.

But I was not only lucky to have such a fantastic primary supervisor, but also a family of associate and co-supervisors – Anne Smith, Amity Campbell, Peter O'Sullivan, and JP Caneiro. I was warned by fellow PhD candidates about the perils of having such a large group of supervisors, but you proved them wrong. This team is like no other. How fortunate I am to have had a team with such a broad range of skills and experience. As a team, you have demonstrated to me how a healthy, co-operative, and constructive team operates – I thank you for this. I will endeavour to foster this approach to teamwork throughout my career.

Thank you to the team at the Curtin Institute for Computation for their support. Nancy Tippaya, Behrouz Beheshti and Kathryn Napier – your collaboration was invaluable to the completion of the body of work presented within this thesis. You were all so friendly and patient in helping me develop my understanding of machine learning. Paul Davey, at one point I thought I'd never get the data I needed to complete my studies. Thank you for the many hours you put in and the friendship you showed while I was stuck in the dungeon. I'm sure I will again need 'A program written by Paul'. Tara Binnie thank you so much for your friendship and assistance in collecting data. I will forever remember the days (weeks) we were suck in the dungeon, punching away at the computers, without aircon, with a familiar rhythmic sound in the background. Eddie Tikoft, thank you so much for all those hours we

spent together milling through the thousands of articles, it was a pleasure working with you.

Thank you to the research participants who volunteered their time, I hope the work in this thesis may help you in the future. To Richard Wright and Rosette Corte-Real your assistance with equipment and facilities was so helpful, thank you for helping me avert potential disasters. And finally, my fellow PhD candidates – the PhatDocs, I'm so glad we had each other during our journeys, you were all so helpful and supportive, I'm so glad I have made such wonderful friends.

# Dedication

To my most wonderful, supportive, and patient wife Judith. What a ride it has been. I couldn't have done this without you! To Amelia, I know you will grow into a strong and independent thinker. I will forever remember how you helped me press the spacebar while sitting on my lap finalising my thesis. I hope this thesis demonstrates that perseverance and ambition can lead to wonderful accomplishments.

# Chapter 1
# Thesis Introduction

## 1.1 Background

Knee osteoarthritis is a growing problem and a leading cause of musculoskeletal disability worldwide (Safiri et al., 2020; Vos et al., 2012). The impact of the condition on the individual is characterised by symptoms including pain, joint stiffness, sleep disturbance, lower limb weakness, and emotional distress (Hawker et al., 2008; Wallis et al., 2019). Symptoms result in disability characterised by activity limitation performing daily activities and participation restriction when performing activities in society (World Health Organization, 2002).

People with knee osteoarthritis most commonly report activity limitation with the activities of walking (Machado et al., 2008; Wallis et al., 2019; Wilkie et al., 2007), negotiating stairs and transitioning to and from a chair (Machado et al., 2008; Wallis et al., 2019; Wilkie et al., 2007). There is consistent evidence in the literature that indicates that people with knee osteoarthritis have different movement patterns when performing those activities compared to people who do not have knee osteoarthritis (Iijima et al., 2018; Mills et al., 2013; Sonoo et al., 2019; van Tunen et al., 2018). Movement patterns are characterised by the differences in specific movement parameters such as knee adduction and flexion moment (Iijima et al., 2018; Mills et al., 2013; Sonoo et al., 2019; van Tunen et al., 2018), knee flexion angle (Iijima et al., 2018; Mills et al., 2013; Sonoo et al., 2019), and hamstring-quadriceps co-contraction (Iijima et al., 2018; Mills et al., 2013). Some of these movement patterns have been found to be associated with structural progression (Chang et al., 2015; Chehab et al., 2014; Hodges et al., 2016), activity limitation and pain (Hall et al., 2017; Nebel et al., 2009; O'Connell et al., 2016).

Exercise is a core guideline recommendation for people with knee osteoarthritis (Bannuru et al., 2019). Some exercise interventions such as neuromuscular exercise

(Ageberg & Roos, 2015) and gait retraining (Richards et al., 2017) target movement patterns directly, while other interventions like progressive resistance training (Foroughi et al., 2011a) or walking programmes (Hunt et al., 2018) do not. Most commonly, the purpose of targeting movement patterns is to change or normalise specific movement parameters that are associated with structural progression, activity limitation or pain. Exercise that aims to change movement patterns have demonstrated efficacy for improving activity limitation and pain in randomised controlled trials (Bennell et al., 2010; Bennell et al., 2014; Hunt et al., 2018). However, the literature is mixed when exploring if movement patterns change, and therefore it is unclear if there is a relationship between changing movement patterns and clinical outcomes of activity limitation and pain after exercise interventions.

In clinical research the gold-standard equipment for assessing movement patterns in people with knee osteoarthritis requires optoelectronic systems for collecting kinematic data, and with the addition of force-plates kinetic data can also be obtained. However, there is some indication in the literature that people do not perform activities in the same way when they are observed (e.g. in a research or clinical environment), compared to when they are unobserved in free-living environments (e.g. at home or work) (Brodie et al., 2017; Brodie et al., 2016; Del Din et al., 2016; Dreischarf et al., 2016; Renggli et al., 2020; Robles-García et al., 2015; Weiss et al., 2011).

Over the past 25 years there has been growing research interest in wearable sensor systems (Picerno, 2017). While there is growing commercially available wearable sensors such as inertial measurement units (IMUs), clinical uptake is slow. There are limitations using IMUs to collect movement-based data in free-living environments. IMU software most commonly uses fusion algorithms to integrate data from onboard hardware that includes triaxial accelerometers, gyroscopes, and magnetometers (Weygers et al., 2020). Calibration is required to mitigate known errors that occur in the gyroscope and magnetometer, and that is a burden for the user. More importantly, electromagnetic interference can reduce the reliability of IMU measurements particularly in free-living environments where it is not possible to

control the magnetic field. It is also not possible to directly record kinetic data in free-living environments using IMUs alone.

There is significant growing interest in advanced computation approaches such as machine learning for handing IMU data. Monitoring of IMU data in free-living environments has the potential to provide clinicians and researchers information about activity avoidance (reduced movement quantity) and altered movement patterns (altered movement quality). An example of activity avoidance is someone using a elevator, rather than stairs, due to knee pain. An example of an altered movement pattern is a 'quadriceps avoidance pattern' in walking (Al-Zahrani & Bakheit, 2002; Fisher et al., 1997; Messier et al., 1992). Machine learning is a branch of artificial intelligence where a model containing specific algorithms is designed to predict a specific outcome based on previous training. IMU data streams collected in free-living environments are long and do not include automated labelling of activities that would provide context about which activity the wearer was performing. One way to automate labelling of IMU data would be through teaching a machine learning algorithm to predict which activities were being performed, a concept known as 'human activity recognition'. There are a significant number of studies that have developed human activity recognition machine learning models, but the majority have investigated only healthy participants (Albert et al., 2012; Arif & Kattan, 2015; Ascioglu & Senol, 2020; Cust et al., 2019; Dobkin, 2013; Fridriksdottir & Bonomi, 2020; Hendry et al., 2020; Qi et al., 2018; Ramanujam et al., 2021; Jindong Wang et al., 2019). Because people with knee osteoarthritis move differently to healthy controls, machine learning models trained on IMU data from healthy participants may lack validity. Previous studies have established poorer prediction accuracy for systems not trained on the intended population (Albert et al., 2012; Emmerzaal et al., 2020; Lonini et al., 2016). Despite growing interest in development of machine learning human activity recognition models, there appears to be no studies investigating this approach in people with knee osteoarthritis.

Other machine learning studies have focused on predicting specific movement parameters from IMU data. While the majority of studies have investigated healthy participants, there are a handful of studies that have investigated prediction of

kinematic and kinetic parameters in people with knee osteoarthritis (He et al., 2019; Renani et al., 2021; Renani et al., 2020; Wang et al., 2020). However, those studies developed models only for walking, and not for other clinically relevant activities such as negotiating stairs and transitioning to and from a chair. These previous models were designed as knee osteoarthritis population-based models, which in theory is generalisable to people matching the inclusion and exclusion criteria. However, none of those studies investigated the effect of personalising models to potentially refine the prediction accuracy.

Implementation of machine learning models for IMU data in clinical practice would require a pipeline-based approach to data handling. Despite this, standalone machine learning models are commonly reported in the literature. Therefore, there are calls for development of IMU data handling pipelines that output data about both human activities and clinically important movement parameters (Kobsar et al., 2020). To date, there is only one study that has developed such a pipeline, although that system was only validated on healthy participants rather than people with knee osteoarthritis (Emmerzaal et al., 2020).

## 1.2 Statement of the Problem

Activity limitation and pain are common in people with knee osteoarthritis and exercise interventions are recommended in clinical guidelines to improve those outcomes. However, it is unclear if changes in movement patterns after exercise interventions are associated with changes in activity limitation or pain across a range of activities, movement parameters and types of exercise. While people with knee osteoarthritis frequently avoid painful activities and exhibit altered movement patterns, clinicians do not have objective means of monitoring the actual performance of activities or movement patterns outside of a clinical or research environment.

## 1.3 Thesis Aims

The broad aims of the research in this doctoral thesis are to:

1. Explore the relationship between a change in movement patterns and change in clinical outcomes following exercise interventions; and
2. Investigate how wearable sensor technology could be used to monitor activity avoidance and altered movement patterns in people with knee osteoarthritis.

## 1.4 Structure of Thesis

This thesis is comprised of seven chapters and a series of appendices.

Chapter 1 introduces the problem of knee osteoarthritis and the current evidence about the relationship between movement patterns and clinical outcomes, limitations of current motion analysis systems and how machine learning that uses IMU data could address those limitations.

Chapter 2 presents a literature review detailing the prevalence and impact of knee osteoarthritis as well as exploring the relationship between common activity limitation, symptoms and movement patterns associated with the condition. Various methods that provide information about physical function are explored. IMU technology is explored in the context of benefits and limitations and how machine learning may help to overcome those limitations.

Chapter 3 contains a systematic review investigating the relationship between a change in knee related movement parameters and change in clinical outcomes (activity limitation or pain) after exercise interventions. This study was published in the *Journal of Orthopaedic & Sports Physical Therapy*.

Chapter 4 reports the development and validation of a machine learning human activity recognition system for people with knee osteoarthritis to classify clinically important activities (walking, negotiating stairs and transitioning to and from a chair) and directions and phases of those activities. This study was published in the journal *Sensors*.

Chapter 5 outlines the development and validation of a machine learning kinematic prediction model for multiple clinically important activities (walking, negotiating stairs and transitioning to and from a chair). This study was published in the journal *Sensors*.

Chapter 6 describes the development and validation of a machine learning kinetic prediction model for the stance phase of walking. This study has been submitted to a journal and is under review.

Chapter 7 is a discussion of the main findings of the thesis. The main findings from the studies within this thesis are explored within the context of individualised assessment and management in clinical practice and research. Opportunities and challenges of implementing machine learning data handling pipelines for IMU data will be described with a focus how the findings from this thesis may inform future clinical practice and research.

# Chapter 2

# Literature Review

This chapter aims to review the scientific literature about (a) the pathology, burden and risk factors of knee osteoarthritis, (b) the relationship between movement patterns and clinical outcomes, (c) interventions for knee osteoarthritis with a focus on interventions that target a change in movement patterns, (d) monitoring outcomes in people with knee osteoarthritis, (e) using inertial measurement units to measure physical function in people with knee osteoarthritis, and (f) using machine learning for processing inertial measurement unit data for human activity recognition and biomechanical analysis.

## 2.1 Knee Osteoarthritis

### 2.1.1 Pathology and Diagnosis of Knee Osteoarthritis

The Osteoarthritis Research Society International (OARSI) defines the pathology of osteoarthritis as:

"…[a] disorder involving movable joints characterised by cell stress and extracellular matrix degradation initiated by micro- and macro-injury that activates maladaptive repair responses including pro-inflammatory pathways of innate immunity. The disease manifests first as a molecular derangement (abnormal joint tissue metabolism) followed by anatomic, and/or physiologic derangements (characterised by cartilage degradation, bone remodelling, osteophyte formation, joint inflammation and loss of normal joint function), that can culminate in illness." (Kraus et al., 2015)

While the definition of osteoarthritis as a pathology is biomedical in nature, the clinical presentation requires consideration of a range of biopsychosocial factors (Caneiro, O'Sullivan, et al., 2020; Hunter, 2018; Kittelson et al., 2014). Knee osteoarthritis arises from a combination of modifiable and non-modifiable risk factors that relate to inflammation and subsequent reduction in cartilage tissue quality

(Mobasheri & Batt, 2016). Traditionally, a diagnosis of knee osteoarthritis was obtained through medical imaging (Kellgren & Lawrence, 1957); however, medical imaging alone does not assist in determining symptomatic pathology. Therefore, multiple guidelines recommend diagnosing knee osteoarthritis through a combination of clinical features that include: (a) >45 years of age, (b) activity-related joint pain and (c) no morning joint stiffness > 30 minutes (National Institute for Health & Care Excellence, 2014; Zhang et al., 2010). Nonetheless, imaging is frequently considered for grading the structural severity of knee osteoarthritis. One of the most widely used grading systems for clinical research is the Kellgren-Lawrence system that grades osteoarthritis from 0 (none) to 4 (severe) (Kellgren & Lawrence, 1957).

### 2.1.2 Burden of Knee Osteoarthritis

#### 2.1.2.1 Impact to Society

In Australia, from the 2017 to 2018 estimates, osteoarthritis affects 2.2 million people, which equates to approximately 9.3% of the population, with increasing prevalence with age (Australian Institute of Health and Welfare, 2021). Similarly, in the United Kingdom for adults over the age of 25 years of age, it is estimated the prevalence of knee osteoarthritis is 10.7% (Swain et al., 2020), and 13.4% in the United States (Cisternas et al., 2016). Knee osteoarthritis accounts for 83% of the total osteoarthritis burden (Vos et al., 2012) and is one of the leading causes of musculoskeletal disability worldwide (Safiri et al., 2020; Vos et al., 2012), with a pooled global prevalence of 16% in people over 15 years of age (Cui et al., 2020).

In Australia, the number of people with of osteoarthritis has been forecast to increase from 2.2 million in 2015, to 3.1 million by 2030, resulting in a projected 39.3% increase in health care costs related to increased disability burden and demand for health care services (Ackerman et al., 2018). A significant component of the increasing health care costs is related to the projected 275% increase in the number of joint replacements performed in Australia by 2030 (Ackerman et al., 2019). The increase in knee surgery is proposed to be unsustainable (Ackerman et al., 2019) and result in significant implications for the healthcare workforce.

#### 2.1.2.2 Impact to the Individual

Knee osteoarthritis is a threat to healthy ageing and is associated with obesity (Silverwood et al., 2015), cardiovascular disease (Calvet et al., 2016) and physical inactivity (Dunlop et al., 2011; Skou et al., 2018). Disability associated with knee osteoarthritis also affects the broader economy and personal finances as those with the condition are at higher risk of work loss compared to those without the condition (Sharif et al., 2016).

Knee osteoarthritis is a 'whole person condition' that is influenced by a combination of biological, psychological and social factors, which interact to influence both the pathology and lived experience of the condition (Caneiro, O'Sullivan, et al., 2020). Common symptoms of knee osteoarthritis include pain, joint stiffness, sleep disturbance, lower limb weakness, and emotional distress, although the individual experience of people who have knee osteoarthritis is variable (Hawker et al., 2008; Wallis et al., 2019). These symptoms result in limitations of physical function (activity limitation) and social roles (participation restriction) (Hawker et al., 2008; Wallis et al., 2019). Activity limitations are *"difficulties an individual may have in executing activities",* while participation restrictions are *"problems an individual may experience in involvement in life situations"* (World Health Organization, 2002). The most common activity limitations in people who have knee osteoarthritis include walking (Machado et al., 2008; Wallis et al., 2019; Wilkie et al., 2007), negotiating stairs and transitioning to and from a chair (Machado et al., 2008; Wallis et al., 2019). Subsequently, these activity limitations can result in participation restrictions that includes difficulty with transportation, socialising, work, and recreational activities (Wallis et al., 2019).

### 2.1.3 Risk Factors for Progression of Knee Osteoarthritis

Part of the solution for the increasing prevalence and costs associated with knee osteoarthritis includes improved ability to assess and manage risk factors associated with progression of structural progression, symptoms, and activity limitation. There is a complex interplay of biopsychosocial risk factors related to progression of structural changes as well as clinical outcomes such as activity limitation and pain. *Structural progression* of knee osteoarthritis has been reported to be related to the following

factors: female sex (Silverwood et al., 2015), age, multisite osteoarthritis, high body mass index (BMI)/obesity (Chapple et al., 2011), previous injury (Poulsen et al., 2019; Silverwood et al., 2015), occupational demands (Canetti et al., 2020), quadriceps weakness (Øiestad et al., 2022), radiographic varus alignment (Brouwer et al., 2007; Chapple et al., 2011), and knee biomechanics during walking (Chang et al., 2015; Chehab et al., 2014).

In contrast, *progression of activity limitation and pain* in people with knee osteoarthritis has been reported to relate to depression (de Rooij et al., 2016; Previtali et al., 2020; Zheng et al., 2021), quadriceps weakness (Culvenor et al., 2017), higher baseline pain, bilateral symptoms (de Rooij et al., 2016), lower education, and comorbidities (Previtali et al., 2020). The salience of particular biopsychosocial risk factors varies between and within individuals over time, underpinning the need for individualised assessment and management.

#### 2.1.3.1 Biomechanical Risk Factors in Knee Osteoarthritis

Biomechanical features such as quadriceps weakness (Culvenor et al., 2017; Øiestad et al., 2022), radiographic alignment (Brouwer et al., 2007; Chapple et al., 2011), and specific movement parameters during walking (Chang et al., 2015; Chehab et al., 2014; Hall et al., 2017; Henriksen et al., 2012; Hodges et al., 2016; Nie et al., 2019) have been implicated as risk factors in people with knee osteoarthritis.

The risk of structural progression of medial compartment knee osteoarthritis is associated with increased medial knee load during walking (Chang et al., 2015; Chehab et al., 2014) and static varus radiographic alignment (Brouwer et al., 2007; Chapple et al., 2011). In a small sample of 16 participants with knee osteoarthritis, Chehab et al. (2014) reported moderate correlations ($R^2 = 0.6$, $p = 0.01$) between medial-to-lateral femoral cartilage thickness ratio and baseline knee adduction moment, knee flexion moment and pain over five years. Similarly, in a prospective longitudinal study of 391 knees (204 people), Chang et al., (2015) reported that larger knee adduction moment and knee adduction moment impulse were associated with reduced cartilage thickness over two years. Longer duration of medial co-contraction of quadriceps and hamstrings has also been implicated in structural progression, which results in

increased medial load during walking, is also predictive of medial tibial cartilage loss (Hodges et al., 2016).

These risk factors are important, because they are potentially modifiable through a range of non-surgical interventions (see section 0), although clinicians are unable to routinely assess three-dimensional biomechanics as part of their clinical practice (see section 2.3.2.2).

### 2.1.4 Movement Patterns in People with Knee Osteoarthritis

People with knee osteoarthritis display movement patterns that differentiate them from those without osteoarthritis during activities such as walking (van Tunen et al., 2018), sit-to-stand (Sonoo et al., 2019; Turcot et al., 2012) and negotiating stairs (Iijima et al., 2018). In clinical biomechanics research, a person's movement pattern (i.e. any possible combination of movement parameters) can be evaluated by measuring kinetic, kinematic and/or muscle activity movement parameters. The methods for conducting biomechanical analysis are discussed in more detail in section 2.3.2.2.

Movement parameters that distinguish people who have knee osteoarthritis from healthy control groups include: (a) increased medial knee joint loading (reported as knee adduction moment) (Heiden et al., 2009; Mills et al., 2013; Sparkes et al., 2019; van Tunen et al., 2018), (b) reduced sagittal plane range of movement (Bouchouras et al., 2015; Heiden et al., 2009; Hinman et al., 2002; Iijima et al., 2018; McCarthy et al., 2013; Mills et al., 2013; Segal et al., 2013) and (c) increased co-contraction of the quadriceps and hamstrings (Heiden et al., 2009; Hodges et al., 2016; Mills et al., 2013). **Table 2-1** provides a summary of studies that have investigated differences in movement parameters between people with knee osteoarthritis compared to healthy controls for a range of clinically relevant activities.

**Table 2-1. Studies investigating differences in movement patterns between people with and without knee osteoarthritis.**

| Movement parameter | Activity | | |
|---|---|---|---|
| | Walking | Negotiating stairs/step | Sit-to-stand/ stand-to-sit |
| Kinetic | | | |
| Knee adduction moment | (Heiden et al., 2009; Sparkes et al., 2019; van Tunen et al., 2018) | (Sparkes et al., 2019) | |
| Knee flexion moment | (Heiden et al., 2009; Mills et al., 2013; Sparkes et al., 2019) | (Iijima et al., 2018) | (Epifanio et al., 2008; Sonoo et al., 2019; Turcot et al., 2012) |
| Kinematic | | | |
| Knee flexion angle | (Heiden et al., 2009; McCarthy et al., 2013; Mills et al., 2013) | (Hinman et al., 2002; Iijima et al., 2018) | (Bouchouras et al., 2015; Segal et al., 2013; Sonoo et al., 2019) |
| Muscle activity | | | |
| Hamstring/ quadriceps co-contraction | (Heiden et al., 2009; Hodges et al., 2016; Mills et al., 2013) | | |
| Hamstring/ quadriceps muscle timing | | (Hinman et al., 2002; Iijima et al., 2018) | |

As discussed in the previous section, some movement patterns are related to risk of structural progression of knee osteoarthritis (Chang et al., 2015; Chehab et al., 2014; Hodges et al., 2016). There is also some evidence that some movement patterns during walking are related to pain severity (Hall et al., 2017; Henriksen et al., 2012; Nie et al., 2019; O'Connell et al., 2016). For example, in a cross-sectional study by (O'Connell et al., 2016), 65 participants were categorised into four different levels of pain severity (none, mild, moderate/severe) using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale. The group with moderate-severe pain demonstrated significantly higher peak knee flexion moment during midstance than those without pain (O'Connell et al., 2016). That same study also demonstrated weak evidence ($p = 0.06$) of a difference between participants with no pain, mild pain, and moderate-severe pain for peak knee flexion angle during weight acceptance. Increased knee flexion was observed for those who are symptomatic when controlling for radiographic severity and gait speed (O'Connell et

al., 2016). Those findings suggest that with a modestly larger sample size, the result may be statistically significant.

Recently it has been reported that there is a stronger link between knee kinematics during walking and clinical outcomes (activity limitation and pain), than between severity of structural changes and clinical outcomes (Bensalma et al., 2022). They conducted a secondary analysis of baseline randomised controlled trial data from 415 participants with knee osteoarthritis. Using canonical correlation analysis (a form of multilevel regression analysis that reports associations between sets of variables) they reported significant moderate to strong *global* correlations between knee kinematics and activity limitation (rho = 0.7), and kinematics and pain (rho = 0.6) explaining > 36% of the variance in the activities of daily living and pain subscale of the Knee Injury and Osteoarthritis Outcome Score (KOOS) (Bensalma et al., 2022). The strongest correlations for *individual* kinematic variables with total KOOS score were in the sagittal plane including flexion at push-off (rho = 0.3) and knee flexion excursion during loading (rho = 0.4), while the weakest correlations were related to frontal plane kinematics (rho = -0.24). Bensalma et al. (2022) also reported weak to moderate correlations between structural severity and activity limitation (rho = 0.2), and structural severity and pain (rho = 0.4) suggesting that kinematic parameters are more strongly related to clinical outcomes than structural changes. Together, those findings suggest that clinical decision making may be better informed by how a person moves rather than the severity of knee osteoarthritis when attempting to facilitate improvement in activity limitation or pain. Based on these cross-sectional results, as greater knee flexion is associated with better function and less pain, targeting increased movement in the sagittal plane may potentially facilitate improvements in clinical outcomes.

The relationship between pain (Hall et al., 2017; Henriksen et al., 2012; Nie et al., 2019), activity limitation (Hall et al., 2017; Nie et al., 2019) and knee adduction moment during walking has been investigated according to underlying structural severity assessed using the Kellgren-Lawrence grading scale (Kellgren & Lawrence, 1957). **Table 2-2** summarises the findings for those studies and demonstrates that across these studies the relationship is not consistent.

**Table 2-2. Relationship between knee adduction moment during walking and activity limitation or pain for different grades of knee osteoarthritis.**

| Studies | Kellgren-Lawrence Grade | | |
|---|---|---|---|
| | Grade 2 | Grade 3 | Grade 4 |
| Pain | | | |
| (Hall et al., 2017) | No association | ↑ KAM Impulse<br>↑ Pain | ↑ KAM Peak<br>↓ Pain |
| (Nie et al., 2019) | ↑ KAM Peak/Impulse<br>↓ Pain | No association | No association |
| (Henriksen et al., 2012) | ↑ KAM Peak/Impulse<br>↓ Pain | ↑ KAM Impulse<br>↑ Pain | ↑ KAM Impulse<br>↑ Pain |
| | | No association<br>KAM Peak and Pain | No association<br>KAM Peak and Pain |
| Physical Function | | | |
| (Hall et al., 2017) | No association | No association | ↑ KAM Peak<br>↑ Function |
| (Nie et al., 2019) | ↑ KAM Peak/ Impulse<br>↓ Function | No association | No association |

KAM = (first peak) knee adduction moment

Because the findings from the studies detailed in **Table 2-2** are mixed, they do not provide certainty about the relationship between knee adduction moment and clinical outcomes assessed through patient-reported outcome measures across knee osteoarthritis grades. Those studies used patient-reported outcome measures that ask about pain and activity limitation in the past 48 hours (e.g. WOMAC) and included people regardless of whether their symptoms were related to physical activity or more specifically to walking, potentially contributing to the mixed results. A different approach to testing the relationship between movement parameters and clinical outcomes is to include participants whose pain specifically increases during or after physical activity. Marriott et al. (2019) conducted a retrospective cross-sectional study of 279 participants with medial knee osteoarthritis that investigated the relationship between movement patterns during walking and increased pain. Participants were included if they experienced increased pain (≥1/10 on the visual analogue scale (VAS)) in one knee, and no increased pain in the other knee after six minutes of walking. Within participants, knee angles and moments were compared between the

knee that increased in pain and the other knee that did not increase in pain, allowing calculation of an odds ratio for increased pain in one knee compared to the other knee. The findings from Marriott et al. (2019) are presented in **Table 2-3**. One interpretation of the findings in **Table 2-3** is that interventions could be selected to target specific movement parameters in a direction that is associated with less pain. However, that has yet to be tested.

**Table 2-3. Relationship between activity-related increase in pain and movement parameters** from Marriott et al. (2019)

| | Movement parameter (peak) | Odds Ratio (Confidence interval) | Interpretation |
|---|---|---|---|
| Significant positive relationship between pain and movement parameter | | | |
| Moments | 1st Adduction | 2.8 (2.0-3.9) | An increase in movement parameter is related to an increase in pain. |
| | 2nd Adduction | 2.4 (1.7-3.2) | |
| | Adduction impulse | 6.6 (3.5-12.6) | |
| | Internal rotation | 7.5 (3.3-17.1) | |
| Angles | Varus | 1.3 (1.2-1.4) | |
| | External rotation | 1.0 (1.0-1.0) | |
| Significant negative relationship between pain and movement parameter | | | |
| Moments | Flexion | 0.5 (0.4-0.6) | An increase in movement parameter is related to a reduction in pain. |
| | Extension | 0.6 (0.4-01) | |
| | External rotation | 0.001 (0.00-0.04) | |
| Angles | Flexion | 0.9 (0.8-0.94) | |
| | Extension | 0.9 (0.9-0.97) | |
| | Internal rotation | 0.9 (0.9-0.97) | |

Overall, the literature is clear that movement patterns differ between people with and without knee osteoarthritis. The relationship between movement patterns and clinical outcomes, such as activity limitation and pain, is beginning to emerge in cross-sectional research, although further studies are required to substantiate these early findings and test if targeting these movement parameters is of clinical value. Studies investigating the relationship between movement patterns and clinical outcomes are limited mostly to the activity of walking, despite other activities also being of clinical importance. There are a range of interventions that have the potential to modify movement patterns and improve clinical outcomes. The effect of some of these interventions will be explored in the following sections.

### 2.1.5 Key Points

*What is known and not known about knee osteoarthritis?*

- Knee osteoarthritis accounts for most of the osteoarthritis burden with forward projections suggesting a significant increase in disability burden and health care costs over the next decade.
- The most common activities associated with activity limitation and pain include walking, negotiating stairs, and transitioning to and from a chair.
- When performing those activities – kinematic, kinetic and/or muscle activity parameters differ compared to healthy controls.
- Some movement parameters are associated with structural progression, activity limitation or pain, although the exact nature of these relationships is still emerging.

## 2.2 Interventions for Knee Osteoarthritis

The core guideline recommendations for treatment of knee osteoarthritis include education, structured land-based exercise programmes or mind-body exercise with or without dietary weight management (Bannuru et al., 2019). Additional management can also include topical, oral, or injectable anti-inflammatory pharmacological agents (Bannuru et al., 2019). It is recommended that these approaches be trialled prior to referral to an orthopaedic surgeon for consideration of a knee arthroplasty (replacement).

There is considerable research investigating the effects of land-based exercise in people with knee osteoarthritis. The most recent Cochrane systematic review reported that there is high quality evidence for improvement in pain and quality of life and moderate quality evidence of improved physical function (reduced activity limitation) for land-based exercise (Fransen et al., 2015). For those outcomes, there are moderate effect sizes immediately post-intervention and small effect sizes at two to six months post-intervention (Fransen et al., 2015). Most commonly, the outcomes are assessed via patient-reported outcome measures which provides information about a person's perception of what they can do, rather than objectively quantifying what they can actually do, a concept that is explored in section 2.3.

Among land-based exercise interventions, there are many different approaches that target various aspects of impairments of movement or fitness. The most common targets include improving lower limb muscular strength, aerobic fitness, range of movement or neuromuscular control (Fransen et al., 2015). Within this thesis, exercise will be defined per the Medical Subject Headings (MeSH) of 'exercise therapy' – "[a] regimen or plan of physical activities designed and prescribed for specific therapeutic goals. Its purpose is to restore normal musculoskeletal function or to reduce pain caused by diseases or injuries" (National Centre for Biotechnology Information, 2021). The most commonly investigated land-based exercise interventions (Bannuru et al., 2019; Fransen et al., 2015) include strength/resistance training, aerobic exercise programmes (e.g. walking or cycling), mind-body interventions (e.g. yoga or Tai-Chi), balance or proprioception interventions (Runhaar

et al., 2015), and neuromuscular exercise (Ageberg & Roos, 2015; Bennell et al., 2014; Holsgaard-Larsen et al., 2017; Skou & Roos, 2017).

There are multiple theories about the mechanism of exercise-induced pain reduction in people who have knee osteoarthritis (Davis et al., 2020; Runhaar et al., 2015). A narrative review by Davis et al. (2020) described general mechanisms, including widespread anti-inflammatory and antinociceptive neural effects, and cartilage and inflammatory biomarkers. Other psychosocial factors have also been implicated, including reduced helplessness, increased self-efficacy, as well as improved health beliefs, coping and mood (Hurley et al., 2003).

A systematic review investigating possible mechanisms of effect for exercise interventions identified 12 potential mediators, although no studies in that review performed a formal mediation analysis (Runhaar et al., 2015). The potential mediators most investigated across all the studies were muscle strength (61 intervention groups), gait properties and biomechanics (25 intervention groups) and range of movement (21 intervention groups) (Runhaar et al., 2015), suggesting these features may be important to target as part of an exercise-based intervention. Guidelines consistently recommend exercise targeting improvement in muscular strength which can help to address deficits in quadriceps strength, identified as a risk factor for progression of knee osteoarthritis outlined in section 2.1.3. However, the role of changing movement patterns remains unclear. Therefore, the ongoing focus of this literature review will explore (a) interventions that target a change movement patterns and (b) how physical function is assessed in people with knee osteoarthritis.

### 2.2.1 Interventions that Target Movement Patterns

Because of the relationship between some movement parameters and structural progression (Chang et al., 2015; Chehab et al., 2014), activity limitation and pain (Hall et al., 2017; Henriksen et al., 2018; Hodges et al., 2016; Marriott et al., 2019; Nie et al., 2019), interventions have been developed that aim to modify movement patterns in order to slow or prevent structural progression, as well as reduce activity limitation and pain. Such interventions include knee support braces, orthotics or prescription footwear (Radzimski et al., 2012), gait retraining (Richards et al., 2017) and neuromuscular exercise (Ageberg & Roos, 2015).

While there are a variety of interventions that target abnormal movement patterns in people with knee osteoarthritis, current guidelines only support exercise (Bannuru et al., 2019; Royal Australian College of General Practitioners, 2018). Other interventions that have the potential to change movement patterns such as orthotics and knee braces are not supported in guidelines, largely due to the low quality of available studies.

### 2.2.2 Strengthening Exercise, Movement Patterns and Clinical Outcomes

Some exercise interventions, like strengthening exercise or resistance training do not target a change in movement patterns directly. Resistance training is typically targeted at strength deficits in the quadriceps muscles, and also has the potential to improve knee confidence (Skou, Rasmussen, Simonsen, et al., 2015), self-efficacy, inflammatory mediators, activity limitation and pain (Runhaar et al., 2015). Improvements in quadriceps function also have the potential to subsequently modify sagittal plane movement (e.g. flexion angle) or loading patterns (e.g. flexion moment) of the knee during functional activities. However, across studies, strengthening exercise does not change early stance phase peak knee flexion angle during walking, despite improvements in activity limitation and pain (DeVita et al., 2018; Fisher et al., 1997; Gaudreault et al., 2011). Similarly, in a cohort study of 14 participants, Al-Khlaifat et al. (2016) demonstrated significant improvements in activity limitation and pain, without concurrent changes in knee adduction moment (peak or impulse) during the stance phase of walking after six-weeks of combined weight bearing resistance training and balance exercise. They did however report a significant reduction in early and mid-stance lateral hamstring-quadriceps co-contraction (Al-Khlaifat et al., 2016). McQuade and de Oliveira (2011) investigated the effect of eight-weeks of progressive machine-based resistance exercises on knee moments and muscle activity during a step-up task. They reported significant improvements in activity limitation and pain, without concurrent changes in knee flexion and extension moments, changes in muscle activation amplitude of the quadriceps and hamstrings, or changes in hamstring-quadriceps ratio. Together, these findings suggest that improvements in clinical outcomes may occur in the absence of changes in movement patterns following land-based resistance training. While it is possible that the most common

forms of assessment of movement parameters may lack sensitivity to detect change, two other possible reasons that movement patterns do not change after strengthening-based exercise is that the intervention does not target movement patterns directly (see section 2.2.3), or that the strength increase was insufficient to result in a change in movement pattern.

While most studies mentioned above do not demonstrate change in movement parameters, Davis et al. (2019) explored the effect on movement parameters after four-weeks of lower limb strengthening exercise on responders compared to non-responders with knee osteoarthritis. Participants were classified as responders when improvement in quadriceps strength improved more than the upper limit of the 95% confidence interval (CI) for the minimal detectable change. Based on a 2 x 2 functional ANOVA which can identify significant differences between groups for time-series data, the authors found significantly greater knee flexion angle during the first 50% of stance and significantly greater knee extension in the second 50% of stance in responders compared to non-responders (Davis et al., 2019). These findings suggest that sagittal plane kinematics are affected by changes in quadriceps strength following exercise in people with knee osteoarthritis. However, these changes in kinematics may be unrelated to changes in pain (WOMAC pain/VAS pain) or function (WOMAC function) as there were no between-group differences in change scores between responder and non-responder groups.

In summary, clinical outcomes may change after exercise because of a diverse range of biopsychosocial factors. Currently, there does not seem to be evidence in the literature that movement patterns change frequently as a result of exercise interventions that do not directly target movement parameters. There may be some indication that for those people who improve in strength, they may also have a concurrent change in movement patterns and that concept needs to be further investigated. However, there are other approaches to exercise that directly target movement patterns.

### 2.2.3 Exercise Approaches that Target Movement Patterns

Exercise interventions targeted at changing movement patterns have been theorised to improve pain and physical function through improved neuromuscular

control of the knee and lower limb (Harris-Hayes et al., 2010; Lehman, 2018). There is no single agreed definition for neuromuscular control. It is generally accepted that neuromuscular control describes the ability to co-ordinate the musculoskeletal system under differing external knee loads and is influenced by strength, proprioception, balance, relative amplitude of contraction of antagonist musculature and passive ligamentous support (Ageberg & Roos, 2015; Mills et al., 2013; van der Esch & Dekker, 2014). Therefore, as neuromuscular control is a multidimensional construct, there are a variety of different kinetic and kinematic movement parameters that have been investigated for change after exercise that targets a change in movement patterns.

Two popular exercise-based interventions targeting movement patterns are neuromuscular exercise and gait retraining. Neuromuscular exercise is included as part of the OARSI guidelines for non-surgical management of knee osteoarthritis due to evidence that it improves activity limitation and pain (Bannuru et al., 2019). In contrast, there is only preliminary data suggesting that gait retraining improves activity limitation and pain (Richards et al., 2017; Richards et al., 2018).

Interventions that aim to directly change movement patterns use verbal, visual or proprioceptive cues to explicitly modify the way in which a movement is performed. The new movement pattern is practiced regularly with the belief that with sufficient practice the habitual movement pattern will be replaced with one that reduces the load on the knee and thereby improving physical function, reducing symptoms and preventing further structural progression of the condition (Ageberg & Roos, 2015).

**Neuromuscular exercise** is a popular intervention for knee osteoarthritis, promoting normal alignment of the lower limb (e.g. reduce knee adduction moment) during traditional functional strengthening exercises, such as bridges, lunges, step-ups and squats (Ageberg & Roos, 2015). There are several randomised controlled trials that have established that neuromuscular exercise (or combined interventions) is superior to education alone to improve activity limitation and pain immediately post-intervention (da Silva et al., 2015; Hurley et al., 2007; Skou, Rasmussen, Laursen, et al., 2015) and for up to one year (Skou, Rasmussen, Laursen, et al., 2015).

However, in other studies, the effect of neuromuscular exercise is not superior when compared to oral analgesics (Holsgaard-Larsen et al., 2017) or quadriceps strengthening exercise (Bennell et al., 2014) for improving activity limitation or pain. Those studies indicate there was no between-group difference in knee adduction moment change scores between neuromuscular exercise and control groups (oral analgesics or quadriceps strengthening exercise). This suggests that improvement in activity limitation and pain may occur via mechanisms that transcend the proposed target for each intervention, that different mechanisms may exist for each intervention, or that changes in knee joint loading may not be a prerequisite for improvement in clinical outcomes.

Regardless of changes in clinical outcomes, some studies do demonstrate within-group changes in movement parameters. In a randomised controlled trial testing the efficacy of neuromuscular exercise compared to pharmacological management of 93 participants with knee osteoarthritis, Holsgaard-Larsen et al. (2017) reported significant small within-group reductions in knee adduction moment (1$^{st}$ peak and impulse) during walking, that were not observed in the oral analgesic control group. Contrary to those findings, an earlier cohort study investigated the effect of neuromuscular exercise for early-stage knee osteoarthritis on knee adduction moment (Thorstensson et al., 2007). No change in knee adduction moment during walking was found, although there was a reduction in peak knee adduction moment during one-leg sit-to-stand in the affected leg but not the unaffected leg (Thorstensson et al., 2007). This suggests that neuromuscular exercise may selectively alter movement patterns for activities that have similar biomechanical requirements to the type of exercise intervention, while having limited effect on movement patterns of activities that have different biomechanical requirements. For example, squats and lunges performed within a neuromuscular exercise intervention are biomechanically similar to one-leg sit-to-stand compared to walking.

In studies investigating neuromuscular exercise described above, the study methods did not take into consideration baseline biomechanics of participants which has the potential to impact the effect of the intervention. One study has explored that concept using data from randomised controlled trials. Bennell et al. (2015) reported

that neuromuscular exercise improved pain for people who demonstrated varus thrust (a surrogate, subjective, visual measure of knee adduction moment) during walking, compared to those who did not have varus thrust, suggesting baseline movement patterns might influence outcome. The authors describe two possible mechanisms (Bennell et al., 2015). Firstly, that because there is a relationship between varus thrust and pain (Lo et al., 2012), neuromuscular exercise may directly reduce thrust and therefore improve pain. Alternatively, they suggest that because neuromuscular exercise focuses on control of movement, especially in the frontal plane, changes in muscle activation patterns or proprioception may be involved in reducing pain. The authors concede that they did not objectively assess varus thrust, nor muscle activity or proprioception, reducing the confidence of those conclusions.

Together these studies suggest that neuromuscular exercise is efficacious for improving clinical outcomes of activity limitation and pain. However, there is conflicting evidence as to whether walking-related movement patterns change in response to the intervention, and there are a limited number of studies that have investigated changes in other activities important to people with knee osteoarthritis (e.g. sit-to-stand and stairs). Underlying baseline movement patterns prior to neuromuscular exercise may influence changes in symptoms, however it is unclear if post-intervention changes in movement patterns are related to improvements in clinical outcomes.

**Gait retraining** is another exercise intervention (according to the MeSH definition – see section 0) that has been investigated for its potential to modify knee adduction moment and improve clinical outcomes. Unlike neuromuscular exercise, gait retraining is prescribed as an intervention that directly reflects the activity assessed via biomechanical analysis – a participant's gait pattern. A recent systematic review summarised the literature about effects of gait retraining changing in knee adduction moment and/or clinical outcomes of activity limitation and pain (Richards et al., 2017). They identified three studies (Hunt & Takacs, 2014; Segal et al., 2015; Shull, Silder, et al., 2013) that demonstrated medium to large within-group effect sizes for activity limitation (standardised mean difference (SMD) 0.55 to 0.85) and pain (SMD 0.55 to 1.16). Three cohort studies (Hunt & Takacs, 2014; Shull, Shultz, et al.,

2013; Shull, Silder, et al., 2013) demonstrated small effect sizes (SMD 0.29 to 0.37) for reducing knee adduction moment immediately after intervention, while the two other studies did not provide data about change in knee adduction moment (Hunt et al., 2014; Segal et al., 2015). Of the studies that reported change in knee adduction moment, there were also concurrent changes in activity limitation (Shull, Silder, et al., 2013) and pain (Hunt & Takacs, 2014; Shull, Silder, et al., 2013), potentially suggesting a relationship between a change in movement parameters and clinical outcomes.

Since the review by Richards et al. (2017), three additional studies investigating gait retraining in people who have knee osteoarthritis have been published reporting longer term follow-up assessments (Cheung et al., 2018; Hunt et al., 2018; Richards et al., 2018). Two studies were randomised controlled trials investigating gait retraining compared to a walking exercise control group (Cheung et al., 2018; Hunt et al., 2018). Those studies demonstrated large between-group and within-group improvements in WOMAC physical function and pain subscales benefiting the gait retraining group (Cheung et al., 2018; Hunt et al., 2018). However, there were conflicting findings for change in first peak knee adduction moment. Cheung et al. (2018) reported a significant between-group difference and within-group change for *first* peak knee adduction moment during walking, while (Hunt et al., 2018) reported no between-group difference or within-group change. Hunt et al. (2018) also reported significant between- and within-group change for *second* peak knee adduction moment and knee adduction moment impulse. Similar to Cheung et al. (2018), in a small cohort study (n = 21), Richards et al. (2018) found that first peak knee adduction moment significantly reduced after six-weeks of toe-in gait retraining. While the findings of those three studies (Cheung et al., 2018; Hunt et al., 2018; Richards et al., 2018) first appear to be conflicting, one explanation may be the differences in the gait retraining approach (Simic et al., 2013). The study by Hunt et al. (2018) prescribed that all participants adopt a toe-out gait modification affecting second peak knee adduction moment. Richards et al. (2018) prescribed a toe-in modification affecting first peak knee adduction moment. Cheung et al. (2018) provided individualised gait modification based on a pre-intervention assessment and targeted a reduction in first peak knee adduction moment via changing individual participants' movement

patterns including adjusting foot progression angle (either toe-in or toe-out), hip adduction or rotation, or trunk sway as required. A toe-in gait modification is theorised to reduce first peak knee adduction moment during the stance phase of walking resulting in the knee joint centre moving towards the midline, resulting in a more lateral centre of pressure (Shull, Shultz, et al., 2013). Whereas a toe-out modification moves the centre of pressure laterally, because the line of action of the ground reaction force moves closer to the stance phase weight bearing knee, resulting in reduced second peak knee adduction moment (Jenkyn et al., 2008).

The study by Cheung et al. (2018) provides first evidence for sustained changes (six months post-intervention) in both knee adduction moment and clinical outcomes. While other studies demonstrated improvements in clinical outcomes up to six months, they did not demonstrate longer term changes in movement parameters (Hunt et al., 2018; Richards et al., 2018). One possible explanation for this is that the intervention was individualised for each participant – a concept that requires further investigation. Together, studies in gait retraining provide preliminary evidence that suggest that specific training cues (toe-in vs toe-out) may selectively alter specific walking related movement parameters in people with knee osteoarthritis. The literature currently suggests that gait retraining interventions positively influence activity limitation and pain despite differences in movement parameter affected (e.g. first vs second peak knee adduction moment).

**Other forms of movement-based mind-body exercise** exist for people with knee osteoarthritis. Zhu et al. (2016) conducted a randomised controlled trial and demonstrated improvements in activity limitation and pain after 24-weeks of Tai Chi compared to wellness education for people with knee osteoarthritis. There was also between- and within-groups increase in knee flexion angle at initial contact and knee flexion range of movement during walking favouring Tai Chi. In contrast, a 12-week yoga programme for a cohort of participants with knee osteoarthritis did not change knee adduction moment during walking despite significant improvements in activity limitation and pain (Brenneman et al., 2015). Another intervention known as Alexander Technique aims to promote neuromuscular control through increased attention and awareness of muscle tension (Preece et al., 2016). A cohort study of 21

participants with knee osteoarthritis demonstrated that hamstring-quadriceps co-contraction significantly reduced after 12-weeks of Alexander Technique intervention, along with improvements in long-term (15-month post-intervention) measures of activity limitation and pain assessed using WOMAC.

Similar to studies investigating general strengthening exercise, gait retraining and neuromuscular exercise – mind-body exercise has demonstrated consistent improvements in activity limitation and pain across randomised controlled trials and cohort studies. However, there are a limited number of studies investigating changes in movement patterns and clinical outcomes using mind-body interventions. Such few studies make it difficult to draw conclusions about whether changes in movement patterns are related to changes in clinical outcomes.

In summary, exercise interventions consistently improve activity limitation and pain for people with knee osteoarthritis (Bannuru et al., 2019). However, there is conflicting evidence that movement patterns change and if so whether they are related to changes in levels of activity limitation and pain. There are a variety of reasons why movement patterns do not change consistently across studies investigating the effect of exercise in people with knee osteoarthritis. There appears to be more consistent changes in walking-related movement patterns for gait retraining interventions, possibly because the method of intervention and assessment are aligned. In contrast, interventions such as neuromuscular retraining do not seem to affect walking-related movement patterns, which could be, at least in part, because the intervention does not specifically target walking itself.

Another reason may be because of the variability in movement patterns across the heterogeneous population who have knee osteoarthritis (Gustafson et al., 2015; Hunt et al., 2010; Thorp et al., 2006). Some people with knee osteoarthritis may have no capacity to change because their movement patterns are considered normal, while others may have fixed movement impairments (e.g. an inability to fully straighten the knee, or reduce a structural varus deformity). For those with altered movement patterns, the way they move may not be related to physical function or pain, and alternatively, may be adaptive (helpful and pain relieving) rather than maladaptive (unhelpful and pain inducing). It also might be that, in the presence of clinical

heterogeneity, 'one-size-fits-all' interventions where the same strategy to change movement pattern is prescribed to the whole group, potentially results in smaller effects than individually tailored interventions based on that person's baseline movement or control impairments.

### 2.2.4 Relationship between a Change in Movement Pattern and Clinical Outcomes

While there are some studies that demonstrate concurrent change in movement patterns and patient-reported clinical outcomes after strength training (Al-Khlaifat et al., 2016; Davis et al., 2019), neuromuscular exercise (Bennell et al., 2015; Holsgaard-Larsen et al., 2017), gait retraining (Hunt & Takacs, 2014; Shull, Silder, et al., 2013), and mind-body exercise (Preece et al., 2016; Zhu et al., 2016), it is not clear if there is a consistent relationship between post-exercise changes in movement patterns and clinical outcomes.

One systematic review exists that explored this topic, and only for knee adduction moment during walking. Ferreira et al. (2015) conducted a systematic review of randomised controlled trials to investigate if there was a relationship between changes in medial joint loading during gait and clinical outcomes after an exercise intervention in people with knee osteoarthritis. They included one high-quality (Bennell et al., 2010) and two low-quality studies (Foroughi et al., 2011; Lim et al., 2008). Because no between-group mean difference was found for knee adduction moment across studies, the authors concluded that changes in activity limitation or pain were not associated with changes in movement pattern.

It is common for group-level data to be reported in randomised controlled trials and systematic reviews. While group-level data can provide an indication that two outcomes change together across the sample population, this does not mean that those participants who responded to the treatment on one outcome were the same as those who responded on the other. On the other hand, individual person-level data can be used in studies that use correlation analysis to assess the relationship between changes in two outcomes. For example, in a cohort study of 21 participants with knee osteoarthritis, Preece et al. (2016) have investigated the relationship between changes in movement patterns during the stance phase of walking and activity limitation or

pain after Alexander Technique intervention. They conducted correlation analyses between change in quadriceps-hamstring co-contraction and WOMAC pain. They reported one significant correlation from four tested, with a positive moderate correlation ($r = 0.45$, $p < 0.05$) for pre-contact medial quadriceps-hamstring co-contraction and pain, indicating that lower levels of co-contraction are associated with less pain. Medial co-contraction is of particular interest, as reduction of this movement parameter would theoretically reduce medial joint loading, a potentially important change for people with medial compartment knee osteoarthritis.

While Ferreira et al. (2015) have provided the first review into the relationship between changes in movement patterns and clinical outcomes, questions remain because of their study design. Firstly, they included only randomised controlled trials. Because the question is not about efficacy, but about relationships between change outcomes between two variables, both cohort studies and randomised controlled trials would be informative for this research question. Both cohort studies and randomised controlled trials provide an opportunity to explore within-group change in two outcomes through the use of correlation analysis or co-occurrence of change between outcomes. The relationship between change in other movement parameters (beyond knee adduction moment) and clinical outcomes has also yet to be systematically reviewed.

### 2.2.5 Key Points

*What is known and not known about interventions for knee osteoarthritis?*

- To improve activity limitation or pain, core guideline recommendations for treatment of knee osteoarthritis include education, exercise with or without weight management, prior to considering surgery.
- In studies investigating the effects of exercise, there does not seem to be consistent changes in movement patterns.
- Some exercise interventions target movement patterns either directly or indirectly.
- The relationship between a change in movement parameters and changes in activity limitation or pain after exercise is currently unclear.

## 2.3 Monitoring Outcomes in Clinical Practice

### 2.3.1 Patient-reported Outcome Measures

Clinical practice guidelines for knee osteoarthritis recommend two clinical outcomes above all else for monitoring the status of a patient over the course, or after treatment – activity limitation and pain (National Institute for Health & Care Excellence, 2014; Royal Australian College of General Practitioners, 2018). There are several validated self-reported questionnaires, such as the WOMAC (Bellamy et al., 1988) or the KOOS (Roos & Lohmander, 2003).

Recommended questionnaire-based outcome measures for people with knee osteoarthritis (Bellamy et al., 1988; Roos & Lohmander, 2003) commonly evaluate symptoms such as pain, as well as other features such as swelling, stiffness and crepitus. Pain is typically evaluated in two ways. The first uses outcome measures such as WOMAC to rate pain severity associated with a specific activity (e.g. during walking or ascending stairs). The second is to rate average or maximum pain on a 0 to 10 VAS or numerical rating scale (Alghadir et al., 2018) across a specified number of days (e.g. two days or past week). The assessment of activity limitation in the WOMAC (physical function subscale) and KOOS (activities of daily living subscale) is rated in terms of the perception of difficulty when performing a range of activities related to locomotion or transitioning between positions (e.g. walking, ascending stairs, getting off the floor) or completing daily activities (e.g. going shopping, performing heavy domestic duties).

While pain can be only evaluated by self-report, physical function is a multidimensional construct consisting of perceptual as well as observable physical characteristics. Because patient-reported outcomes for physical function can only evaluate what a person perceives they can do, it is also recommended that objective physical outcome measures of what a person can actually do are also evaluated (Dobson et al., 2013). Consideration of both objective physical outcome measures alongside patient-reported outcome measures is believed to provide a more comprehensive understanding of physical function and activity limitation (Dobson et al., 2013; Stratford et al., 2003; Stratford & Kennedy, 2006; Terwee, Mokkink, et al., 2006). This is particularly important as there are multiple studies that clearly

demonstrate no, or minimal relationship between changes in patient-reported outcome measures of physical function and objective physical outcome measures such as performance-based tests (Stevens-Lapsley et al., 2011; Stratford & Kennedy, 2006; Terwee, van der Slikke, et al., 2006) or physical activity monitoring (Verlaan et al., 2015).

### 2.3.2 Objective Measures of Physical Function

There are three methods currently available for the assessment of physical function that assess the actual performance of activities in people with knee osteoarthritis: performance-based tests, biomechanical assessment, and physical activity monitoring.

#### 2.3.2.1 Performance-based Tests

Performance-based tests are assessed by an observer, typically a clinician or researcher who evaluates physical performance by counting the number of repetitions, timing the duration, or measuring the distance of a functional activity (e.g. walking) or combination of functional activities (e.g. timed up-and-go test) (Dobson et al., 2012; Dobson et al., 2013). While performance tests have not yet been comprehensively validated (Dobson et al., 2012), they are recommended for monitoring a patient's progress (Dobson et al., 2013). However, they are arguably not representative of how a person performs an activity during daily life. For instance, the 30-second chair stand test (Jones et al., 1999) has a patient perform sit-to-stand-to-sit as many times as possible in 30 seconds, which arguably would not represent how a person performs sit-to-stand during their daily life. Similarly, another recommended test, the 40-meter fast-paced walk test (Wright et al., 2011), does not provide information about how a person usually walks. Therefore, while performance tests may capture improvement in one dimension of physical function, those tests provide limited information about physical capacity outside of the clinical or research environment. Recently, it has been recommended that performance-based tests should not be routinely used in clinical practice because only the 40-meter fast-pace walk test was found to be responsive and neither that test nor the 30-second chair stand test were found to have sufficient construct validity (Tolk et al., 2019). However, this study validated performance tests against patient-reported outcome

measures. Considering performance tests and patient-reported outcomes are now known to represent different constructs of physical function (Stevens-Lapsley et al., 2011; Stratford & Kennedy, 2006; Terwee, van der Slikke, et al., 2006), this method of validation could be considered questionable. As performance-based tests do not represent true performance of an activity that is usually performed during unobserved day-to-day life, this creates a dilemma for clinicians. Therefore, clinicians currently have limited ability to access objective, data driven methods of monitoring improvement in their patient's everyday performance of activities. To assess physical function, clinicians must currently rely on direct observation of their patients performing activities during a consultation or alternatively via patient-reported outcome measures.

#### 2.3.2.2 Biomechanical Assessment

Human biomechanical analysis is a technologically-assisted physical assessment that evaluates specific movement parameters of kinematics, kinetics, and muscle activity for the purposes of describing observable movement patterns (Bartlett, 2007). Kinematics refers to geometry of movement without reference to what causes the movement (e.g. joint angles and angular velocity), while kinetics refers to the forces that act on the body to create movement (e.g. force and moments) (Bartlett, 2007).

The gold-standard, surface-based method of estimating kinematics is through the use of an optoelectronic motion analysis system such as Vicon (Oxford Metrics Inc., Oxford, UK) (Cuesta-Vargas et al., 2010; Vicon, 2021). This system uses a minimum of three infrared cameras to track retroreflective markers placed on the research participant within a fixed physical space, usually within a laboratory. Three-dimensional locations of each marker are modelled on to a simulated body comprised of rigid segments allowing estimation of each segment position. From this model joint angles can be calculated. With the addition of force plate data, using inverse dynamics (Camomilla et al., 2017), kinetic movement parameters can be calculated, such as a joint moment, which is the rotational force around the joint axis. Muscle activity, on the other hand, requires the use of electromyography equipment that records the electrical activity of muscle tissue (Mills et al., 2013). Readings from electromyography are often expressed in terms of amplitude of muscular activation or

a ratio between antagonist muscle groups as a description of muscular control around the joint.

There is a substantial body of work using biomechanical analysis to investigate the clinical relevance of movement patterns in people who have knee osteoarthritis (see sections 2.1.3.1 and 2.2.1). Despite this, biomechanical assessment is not currently recommended as part of routine clinical practice to assess the outcome of an intervention. Motion analysis laboratories incur a high cost, have space requirements, and require technological expertise to set up and run, preventing access for most clinicians. However, there are systems that exist for assessing movement parameters in clinical practice that do not provide the level of detail of gold-standard systems.

To measure kinematics in clinical practice, camera-based smart phone applications are available (Krause et al., 2015; Milanese et al., 2014). For example, Coach's Eye (TechSmith Corp, Michigan) is a camera-based smartphone application that has demonstrated excellent reliability for measuring range of movement during functional movements such as a squat (intraclass correlation coefficient 0.98, 95% CI 0.96 to 0.99). Concurrent validity of single plane motion analysis of the knee compared to the reference system Vicon was acceptable (mean difference 5°, 95% limits of agreement ranging −17.6° to 7.6°), and a minimum detectable change of 6° (Krause et al., 2015). So, while camera-based smart phone approaches may help quantify kinematics in a clinical environment, there are some limitations. Such systems are limited as they cannot provide information about three-dimensional kinematics, are unable to provide information about kinetics, and do not provide the opportunity to capture data about how a person performs activities in free-living environments.

Assessment of kinetic movement parameters such as knee joint moments and forces require pressure sensing technology like force plates or shoe insoles combined with motion capture (e.g. Vicon) (Camomilla et al., 2017; DeBerardinis et al., 2018). Clinicians do not commonly have access to combined motion capture and pressure sensing technology as they are typically expensive, would take up clinic space, and require professional expertise to install. Insoles provide a cheaper alternative and have the potential to be used in free-living environments, but do not directly measure

force, resulting in imprecise outputs (Chen et al., 2022). A review by Chen et al. (2022) described multiple limitations of pressure sensing technology including humidity, size, heat, electromagnetic interference, device complexity, calibration methods, sampling requirements and energy consumption. If a patient would like to wear a range of shoes, they would be required to move the insoles between shoes, increasing patient burden. While insoles can provide data about modelling kinetic parameters (Oubre et al., 2021) and assist with labelling gait events, they do not provide spatial data required for kinematic motion analysis (Chen et al., 2022). Recently kinematic estimates from wearable sensors known as inertial measurement units (IMUs) have been used to calculate knee moments and forces using inverse dynamics (Karatsidis et al., 2019). However, there are a number of limitations using IMUs in free-living environments described in section 2.4 that may affect the reliability and validity of that approach.

IMUs are becoming widely accessible, due to reducing hardware and software costs, and many require minimal technical expertise to operate. IMUs allow the wearer freedom of movement through space without being confined to a field of view and data from IMUs can be used to provide both kinematic and kinetic data (see sections 2.4.1.1 and 2.5.6). Compared to performance-based tests, IMUs can record how a person naturally performs activities and are therefore one potential technology that could be used to monitor patient outcomes both within the clinical and free-living environments. Together, patient-reported outcome measures, performance-based tests and IMU monitoring represent different aspects of physical function, providing a clinician a more comprehensive picture of their patient's ability to perform activities.

#### 2.3.2.3 Monitoring Physical Activity beyond a Clinical Environment

Monitoring of physical activity requires an electronic system that can record and store movement-based data. Researchers first investigated accelerometer-based physical activity monitoring in the early 1980s (Montoye et al., 1983) for the purposes of quantifying physical activity intensity (energy expenditure) in free-living environments. Since the early 2000s accelerometer-based activity monitors have been available to the general public in the form of smartphones and smartwatches for

tracking energy expenditure through monitoring physical activity intensity or step count, as well as monitoring sleep patterns (Henriksen et al., 2018). Where the aim is to increase physical activity, energy expenditure can be monitored using a simple accelerometer or pedometer. Over time, hardware costs have reduced resulting in wide spread availability of wrist worn sensors for measuring energy expenditure (e.g. Apple watch www.apple.com/watch). Some devices that monitor energy expenditure can also track or monitor activities but those systems are not designed to automatically identify activities, and instead requires the user to select the activity (Apple Inc, 2022). In comparison, a gold-standard, research grade, accelerometer-based device (e.g. activPAL www.palt.com) has the added benefit of automating the recognition of activities and can provide data about the amount of time spent sitting, standing, stepping, and lying (Carpenter et al., 2021) – a function known as 'human activity recognition' (see section 2.5.5). Therefore, the activPAL can provide information about the duration these activities are performed. However, that system is not able to discriminate walking from other ambulatory activities, such as ascending or descending stairs, nor transitions between positions such as sit-to-stand which are important for people with knee osteoarthritis. In addition to not providing a broader capacity for human activity recognition, accelerometer-based devices are unable to provide other potentially important biomechanical information like kinematics or kinetics.

Evidence across healthy and clinical populations suggests that how people move is different when observed in a laboratory compared to real-world environments (Brodie et al., 2017; Brodie et al., 2016; Del Din et al., 2016; Dreischarf et al., 2016; Renggli et al., 2020; Robles-García et al., 2015; Weiss et al., 2011). While those studies did not directly investigate people with knee osteoarthritis, some studies may still be relevant to this clinical population. For example, in older people (>75years), there is evidence of slower gait velocity, lower cadence and higher step time variability in an observed laboratory environment compared to free-living environments (Brodie et al., 2016; Renggli et al., 2020). So far, these studies that compare observed and unobserved movement patterns have reported only spatiotemporal kinematic data, and not angular kinematics, joint moments, or muscle activity. Nonetheless,

considering changes in spatiotemporal parameters are related to changes in angular kinematics and knee joint loading (Zeni & Higginson, 2009), monitoring changes in movement patterns during performance of activities across multiple environments may provide information that is unable to be obtained when a patient is observed during a consultation.

If the aim of management is to reduce activity avoidance by focusing on increasing the amount of walking and stair use for a person with knee osteoarthritis, a system that can monitor daily performance of these activities is needed. Such as system would require a function that can discriminate between types of activities automatically. While some systems are designed to be able to detect body positions like lying, sitting, standing, and walking (Valkenet & Veenhof, 2019) – none are yet capable of identifying a range of ambulatory activities (e.g. negotiating stairs) or transitions between positions (e.g. sit to stand) that are important to people with knee osteoarthritis. Currently, there is no wearable sensor system that is widely used that can reliably be used to monitor a change in the frequency or time spent performing clinically relevant activities such as walking, negotiating stairs, and transitioning to or from a chair for people with knee osteoarthritis over the course of, or after, treatment. Further, no current technology can provide contextualised angular kinematic and kinetic information (see section 2.4.2), that provides details about which activities were performed while their patient is unobserved in free-living environments, for example – when at home or work.

### 2.3.3 Key Points

*What is known and not known about assessing physical function?*

- Physical function can be assessed using patient-reported outcome measures, performance-based tests, biomechanical assessment or monitoring of energy expenditure.
- When physical function is observed by a clinician or researcher, how a person moves may not be a true representation of how an activity is performed in free-living environments.
- Current methods for assessment of physical function do not provide objective data about how frequently or how long a person with knee osteoarthritis performs activities recommended for assessment in guidelines for physical assessment of people with knee osteoarthritis.
- Accelerometer based wearable sensors provide information about body position but not activities associated with activity limitation in people with knee osteoarthritis, nor provide biomechanical data.
- IMUs have potential to monitor people in free-living conditions but have limitations.

## 2.4 Inertial Measurement Units

IMUs are a wearable sensor technology that record movement-based information while the wearer can move around freely in their environment, which is ideal for monitoring patients in free-living environments. IMUs are worn on the body, either adhered to the skin with double-sided, hypoallergenic tape or via elasticated straps. The IMU is placed onto the body part, which is referred to as a segment (e.g. the thigh or shank). Each IMU houses hardware including a triaxial accelerometer, gyroscope, and magnetometer. The accelerometer records linear acceleration, gyroscope records angular velocity, while the magnetometer records the magnetic field strength. For a segment, raw data from each piece of hardware is recorded, but is clinically meaningless, and therefore is typically converted into interpretable data through fusion algorithms (Picerno, 2017). IMU fusion algorithms combine data from the accelerometer, gyroscope, and magnetometer of multiple IMUs. Most sensors systems use proprietary fusion algorithms to estimate segment orientation or joint position. The position and orientation are estimated on the integration of the accelerometer and gyroscope data, while the magnetometer helps stabilise the frontal plane by providing a heading and corrects for drift that occurs through the integration of accelerometer and gyroscope data (Picerno, 2017). These systems, that can provide potentially important biomechanical information, are becoming increasingly accessible for clinicians due to reducing hardware and software costs.

### 2.4.1 Clinical Utility of Inertial Measurement Units

Kobsar et al. (2020) proposed that IMUs could be used to monitor changes in movement patterns outside of a clinical environment using a pipeline of data handling approaches that include human activity recognition followed by biomechanical analysis. IMUs have the potential to be used to assess and monitor changes in movement patterns over time, in both clinical and free-living environments. Tracking changes in the duration and or the frequency of performance of clinically relevant activities such as walking, negotiating stairs, and transitioning to or from a chair could provide clinicians objective information about activity avoidance behaviours and inform clinical decision making.

#### 2.4.1.1 IMUs for Biomechanical Analysis

Most software for these commercially available IMU systems estimate lower limb segment orientation and joint kinematics using fusion algorithms (Picerno, 2017). IMU fusion algorithms to estimate lower limb kinematics during functional activities (e.g. walking, sit-to-stand, squatting, negotiating stairs, running and cycling) have been validated in comparison to gold-standard motion analysis systems such as Vicon (Weygers et al., 2020). That systematic review reported the range of the root mean square error (RMSE) (standard deviation (SD)) for knee joint sagittal plane estimations to be 1.01° (0.11°) to 11.22° (1.09°) across 23 studies. The vast majority of these studies validated IMUs for walking (20/31 studies) with the next most common activities being sit-to-stand and squatting (6/30 studies) (Weygers et al., 2020). Most of those studies included only participants without pathology. In fact, across two systematic reviews and one scoping review up to 2020, there was only one study that validated an IMU for kinematic features in people with knee osteoarthritis (Hafer et al., 2020). Only one other study has more recently investigated the concurrent validity of IMUs against Vicon motion analysis for estimating knee kinematics in people with knee osteoarthritis (Binnie et al., 2021). These two studies compared IMUs proprietary fusion algorithms that presumably use magnetometer data, against an optoelectronic motion analysis system using ZXY Euler angle decomposition. Hafer et al. (2020) reported a RMSE of 0.29° to 0.92° for total sagittal plane knee range of movement during a gait cycle. Binnie et al. (2021) investigated the concurrent validity for peak and time-series kinematic estimations of an IMU against Vicon for multiple clinically important activities that included phases of walking, negotiating stairs, sit-to-stand, and step up/down. They reported RMSE ranging from 1.97° (1.48°) to 3.02° (2.56°) for peak estimations, and 3.72° (3.63°) to 4.67° (4.12°) for time-series estimations. Both these studies in people with knee osteoarthritis reported RMSE within the range of that found in healthy participants, albeit with higher variability, demonstrating potential for their use in this clinical population.

Despite the low RMSE between the IMU and Vicon systems, the limits of agreement reported by Binnie et al. (2021) indicated a large variability in sagittal plane angular estimates with the limits of agreement for peak predictions ranging between –

9.63° to 9.32° and -16.91° to 20.94°. High variability may preclude clinical use because of unreliable estimates, which may be related to differences in movement patterns in people with knee osteoarthritis or electromagnetic interference (see section 2.4.2) affecting the fusion algorithm. Because no other study has reported variability in the estimates for IMU systems in people with knee osteoarthritis, it is unclear if this is consistent across different IMU systems, knee osteoarthritis populations and movement patterns.

The range of walking-related movement parameters investigated using IMUs for people with osteoarthritis have been summarised in a recent scoping review (Kobsar et al., 2020). From 72 included studies: 45 investigated spatiotemporal parameters, 33 investigated joint angles or segment orientation, 22 investigated accelerations, 10 investigated side-to-side symmetry of walking parameters, and three investigated knee joint moments. They identified 10 studies that investigated knee joint angles and three studies that investigated knee joint moments using IMUs in people with knee osteoarthritis (He et al., 2019; van den Noort et al., 2013; Wang et al., 2020). Kinematic outcomes in those studies included: sagittal plane hip, knee and ankle joint as well as lower limb segment (thigh, shank or foot) range of movement, and other information such as stride duration and step length (van der Straaten et al., 2018). Some systems can also segment data to provide additional clinically important information, such as the kinematics involved during specific phases of an activity (e.g. stance and swing during walking) (van der Straaten et al., 2018). These studies have demonstrated potential for providing clinicians and researchers movement-based information that may be clinically important. However, studies that use IMUs to collect movement-based information to inform clinical reasoning have only recently been investigated.

While there is limited evidence of group-based changes in movement patterns following exercise interventions (see section 2.2.3), there is recent evidence that IMUs can be used to monitor patient-specific movement patterns in people with knee osteoarthritis (Kobsar & Ferber, 2018) and low back pain (Wernli, O'Sullivan, et al., 2020). For example, using IMUs, Kobsar and Ferber (2018) investigated if patient-specific changes in treadmill walking-related movement patterns were

associated with changes in patient-reported outcome measures (KOOS) following a progressive strengthening and balance based exercise intervention for people with knee osteoarthritis. Using a machine learning approach (see section 0), they found a strong association between changes in individualised walking-related movement patterns and clinical outcomes (Spearman's rho = 0.78) (Kobsar & Ferber, 2018). Their approach used principal component analysis to reduce linear acceleration (accelerometer) and angular velocity (gyroscope) data – representing 95% of the total variance of the collected data. The authors justified this approach suggesting that univariate approaches to biomechanical analysis (typically knee adduction moment) in group-based study designs (e.g. randomised controlled trials) have questionable sensitivity across a heterogeneous population like those with knee osteoarthritis. Because the approach was individualised, and used principal component analysis, it is unclear which biomechanical variables were of clinical interest. So, while this study demonstrated change in movement can be assessed using IMUs following exercise intervention, it is unclear how a clinician would practically target a change in movement pattern because the specific kinematic variables are unknown.

More recently, IMUs have been used to monitor changes in individualised lumbar spine movement patterns within a clinical environment for people with low back pain within an experimental single case design of 12 participants (Wernli, O'Sullivan, et al., 2020). The authors reported that 10/12 participants had significant strong to excellent correlations between changes in individualised movement parameters and clinical outcomes. The most common changes were an increase in lumbar range of movement and speed of movement into flexion which was interpreted as being 'less protective' (Wernli, O'Sullivan, et al., 2020). While no studies have used IMUs and described the exact movement parameters of a 'less protective' movement pattern in knee osteoarthritis, they may be similar to people with low back pain. For example, in the previously described (see section 2.1.4) study by Marriott et al. (2019), an increase knee flexion during walking was associated with less pain – potentially suggesting a 'less protective' movement pattern.

While these two studies included only small sample sizes, these studies suggest that IMUs have the potential to be useful in clinical practice to monitor changes in

individualised movement patterns in clinical environments. It is unclear however if these changes in movement patterns would also be evident outside of the clinic in a free-living environment.

#### 2.4.1.2 IMUs for Recognising Activities

Different types of activities have distinct, recognisable movement patterns. Because of these distinct movement patterns, models can be created to detect specific activities from IMU data, an approach known as 'human activity recognition' (Kim et al., 2010). Human activity recognition is an automated approach that has the potential to unobtrusively monitor the performance of functional activities (i.e. activity avoidance) outside of a clinical environment, described in this thesis as 'free-living environments' (see section 2.5.5).

IMU-based human activity recognition has been mostly investigated in people without health conditions, where data is automatically handled by classification models (see sections 2.5.4and 2.5.5) to predict common every day activities such as walking, running, ironing, vacuuming, opening a refrigerator and drinking from a cup (Arif & Kattan, 2015; Ascioglu & Senol, 2020). There are multiple systematic reviews describing the various uses for human activity recognition. IMU-based human activity recognition has been developed for detecting sports-specific bodily movements to monitor training load (McGrath et al., 2020), exercise detection (O'Reilly et al., 2018), and for quantifying activities in people with mobility impairments (Rast & Labruyère, 2020) that could help inform clinical decision making.

However, despite sensor based human activity recognition having potential to recognise clinically important activities for the purpose of informing clinical decision making, there are a limited number of studies involving people with knee osteoarthritis, or related conditions. A review by Rast and Labruyère (2020) identified only one study that recruited participants with knee osteoarthritis (Verlaan et al., 2015), and two others in participants who had total joint arthroplasty (joint unspecified) (Lipperts et al., 2017), and rheumatoid arthritis (Andreu-Perez et al., 2017). In studies by Verlaan et al. (2015) and Lipperts et al. (2017), IMUs were used to identify lying down (duration), sitting (duration), standing (duration), walking (duration/number of steps) and stair climbing (duration/number of steps). In

comparison Andreu-Perez et al. (2017) used IMUs to recognise changes in body position, such as lying to sitting, lying to standing and sitting to standing/walking. While the purpose of the studies by Lipperts et al. (2017) and Andreu-Perez et al. (2017) was to develop and validate human activity recognition models, Verlaan et al. (2015) compared activity profiles between participants with and without knee osteoarthritis without describing validation of their model. Therefore, no study has yet validated a human activity recognition model for people with knee osteoarthritis. Validation is an important step in development of human activity recognition and is explored in section 2.5.3.

### 2.4.2 Limitations of Inertial Measurement Units

While there is promising research investigating the use of IMUs for clinical purposes, there are several limitations that preclude widespread clinical use. As discussed in the previous section, most IMU systems that provide interpretable biomechanical data have only been validated for use in healthy participants in laboratory conditions. Only a few studies have validated IMU systems for people with knee osteoarthritis for kinematic parameters (Binnie et al., 2021; Hafer et al., 2020) and joint moments (He et al., 2019; van den Noort et al., 2013; Wang et al., 2020). No systems have been validated for human activity recognition in people with knee osteoarthritis (Kobsar et al., 2020) which is described further in section 2.5.5.

All IMUs are prone to integration drift (position or orientation errors that grow with time) over longer time periods, such as hours of use (Kok et al., 2017). Drift originating from the gyroscope reflects a constant bias and measurement noise resulting in an accumulated error, and can only be imperfectly corrected via integration of orientation data from the magnetometer (i.e. heading or tilt estimates from the accelerometer for the vertical plane) (Kok et al., 2017; Weygers et al., 2020). These drift corrections are usually based on joint constraint or degrees of freedom boundaries (Weygers et al., 2020). Because of the requirement to correct for drift, both magnetometer and accelerometer-based drift correction methods are limitations of IMU systems.

Because IMU systems are usually validated in controlled laboratory environments, it is unclear if kinematic estimates in clinical or free-living

environments are equally as precise. This is particularly important as the magnetometer in the IMU is susceptible to electromagnetic interference and magnetisation from ferrous materials and electrical systems (Bachmann et al., 2004; de Vries et al., 2009; Schall et al., 2016). Considering the widespread use of ferrous material for construction (e.g. door frames) and furniture (e.g. chairs), as well as electrical equipment such as lighting, wiring, mobile phones and laptops (Bachmann et al., 2004), it is not possible to avoid these sources of interference in both clinical and free-living environments. This makes IMU systems that use magnetometers unreliable in uncontrolled environments resulting in some researchers discarding the magnetometer data because kinematic estimates from fusion algorithms were unreliable as a result of electromagnetic interference (Schall et al., 2016).

To address the issue of electromagnetic interference, some fusion algorithms have been developed that only require data from accelerometers and gyroscopes, using short term accelerometer-based drift correction (Schall et al., 2016; Teufl et al., 2019; Weygers et al., 2020). However, these systems require calibration procedures to be frequently performed and the IMUs to be placed in the same orientation, which places increased technical burden on the user, especially in free-living environments if a patient is required to initiate the IMUs themselves.

While magnetometer-free IMUs are well suited for field use over longer periods of time (Weygers et al., 2020), they create large files of long continuous streams of unlabelled data. Continuous datasets that include hours of information are meaningless for a clinician or researcher without data labelling that describes which activity was being performed during the data collection period. Because of this, significant data handling is required to extract meaningful samples from these data streams. To overcome these barriers, some systems allow the user – either the clinician or patient to timestamp an event manually to create a data collection 'window', which is cumbersome and time-consuming. A more user-friendly option would be to use a human activity recognition approach to automate segmentation and label the data into clinically relevant samples from these large, continuous data files. Another limitation when using IMUs is a lack of certainty about the number required to be worn that would provide accurate results. For example, for human activity

recognition, when there are multiple layers of classification, accuracy can be improved by as much as 20% by using five IMUs compared to one IMU (Hendry et al., 2020). However, for biomechanical analysis in people with knee osteoarthritis the accuracy of biomechanical models has not been tested using different numbers of IMUs and locations.

One approach to help automate handling of large, complex datasets, from IMUs, is known as machine learning, a form of artificial intelligence. Machine learning offers a data driven method to overcome the issues with electromagnetic interference, calibration requirements and data segmentation limitations of inertial measurement units.

### 2.4.3 Key Points

*What is known and not known about IMUs?*

- IMUs are a wearable sensor technology capable of collecting data in free-living environments.
- IMUs can provide kinematic information using fusion algorithms.
- IMUs can not provide kinetic information directly.
- Preliminary studies indicate IMUs can provide individual person-level data that is clinically relevant.
- IMU data can be used for human activity recognition.
- Fusion algorithms commonly rely on a magnetometer to provide meaningful information.
- There are limitations using IMU systems based on fusion algorithms in free-living environments because of electromagnetic interference and calibration requirements.
- The effect of different numbers of IMUs for human activity recognition and biomechanical analysis is unclear in people with knee osteoarthritis.
- Machine learning may help address limitations of current IMU systems.

## 2.5 Artificial Intelligence

### 2.5.1 Background of Artificial Intelligence and Machine Learning

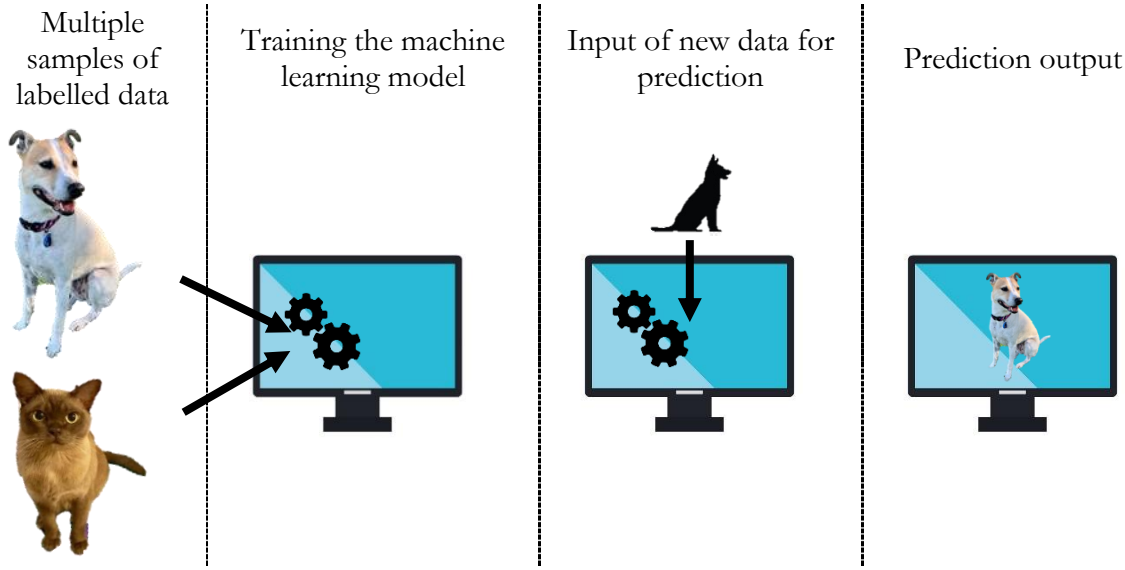
In 1950, the famous mathematician and computer scientist Alan Turing published a seminal paper, posing the question "Can machines think?" (Turing, 1950). He laid the framework for developing and testing artificial intelligence, a concept where computers or machine systems are capable of performing tasks that would usually require human intelligence (Oxford English Dictionary). Machine learning is a specific branch of artificial intelligence based around the idea that computers can learn through experience (Mitchell, 1997; Samuel, 1959). With the correct data and sufficient experience, a machine learning model can be trained to accurately predict an outcome.

There are three types of machine learning – supervised learning, semi-supervised learning, and unsupervised learning (Bi et al., 2019; Wiemken & Kelley, 2020). At its most basic level, unsupervised learning uses unlabelled data for identifying hidden patterns or clusters in large datasets without a pre-specified outcome. On the other hand, supervised learning uses labelled data with a predetermined output which can be used for classification and regression problems.

Classification is most easily understood as a categorical decision-making process. Consider the images of a 'dog' or 'cat' (**Figure 2-1**). A machine (computer algorithm), much like a child, can begin to correctly identify (classify) a furry animal as a dog if presented with sufficient examples of a dog (experience). If the machine is then presented with a cat, it would likely classify it as a dog if it has not been provided with experience to learn the relevant features that provide sufficient information to classify the cat appropriately.

Regression problems on the other hand, can be solved using a supervised prediction model for continuous outcomes. One such example is weather forecasting to predict hourly temperature or rainfall where previous time-series data is used to forecast or predict future events.

**Figure 2-1. Example of a machine learning classifier for cats and dogs.**

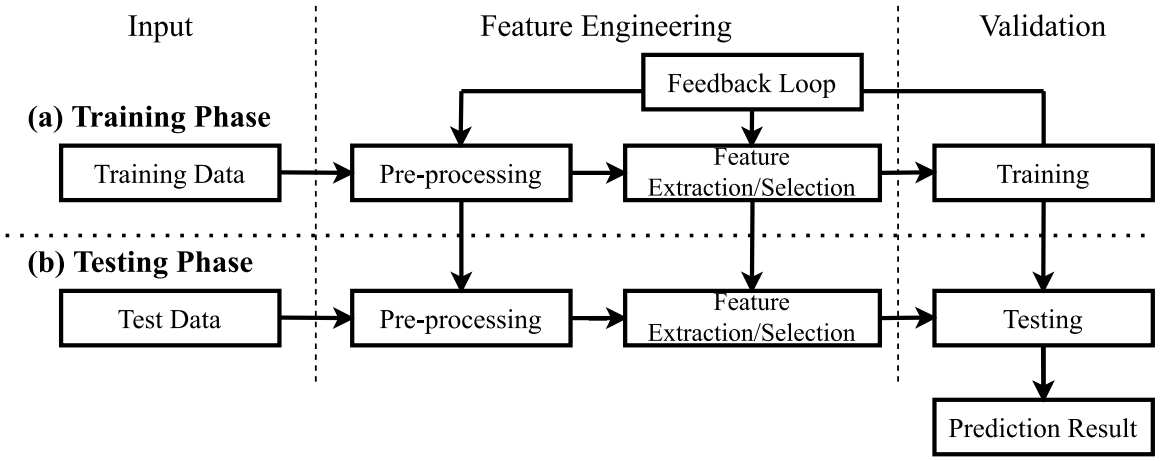



#### 2.5.1.1 Developing Machine Learning Systems

The development of a machine learning system (**Figure 2-2**) broadly involves two phases: (a) training (learning) and (b) testing (prediction) (Kokkotis et al., 2020). The training phase involves an iterative feedback loop that informs all steps of the test phase. Before data is input into the machine learning model, it is important that the data undergoes pre-processing to be 'cleaned'. For this cleaning process, data is checked for consistency, missing samples and any noise is filtered (Kokkotis et al., 2020). After cleaning, further pre-processing can include data transformation (e.g. normalisation) and restructuring. The input data consists of two categories; the target variable (reference standard) and the predictor variable (input data mapped to the target) (Kokkotis et al., 2020). In statistical terms, the target variable is the dependent variable, and the predictor variable is the independent variable.

After the pre-processing phase, the data then undergoes a feature extraction phase where important features are selected that can be used for training the model. This feature extraction phase is part of a feedback loop that provides multiple levels of refinement based on the results of training the model (Kokkotis et al., 2020).

**Figure 2-2. Traditional machine learning development and testing.**




Note. "A typical machine learning system" Adapted from " Machine Learning in Knee Osteoarthritis: A Review" by C. Kokkotis, S. Moustakidis, E. Papageorgiou, G. Giakas and D.E. Tsaopoulos, 2020, Osteoarthritis and Cartilage Open, 2(23), p. 2. (https://doi.org/10.1016/j.ocarto.2020.100069). Copyright 2020 by Osteoarthritis Research Society International (Creative Commons).

#### 2.5.1.1.1 Traditional Machine Learning

There are a variety of traditional machine learning approaches that can handle continuous or categorical input data. The various traditional machine learning approaches can be categorised as Bayesian, linear regression, tree-based, instance-based and support vector machines (Kokkotis et al., 2020). Some examples of traditional machine learning architecture are listed in **Table 2-4**.

**Table 2-4. Types of traditional machine learning approaches.**

| Category | Models | Data |
|---|---|---|
| Bayesian | Naïve Bayes<br>Gaussian Naïve Bayes<br>Multinomial Naïve Bayes<br>Bayesian Belief Network | Categorical |
| Instance-based | K-nearest neighbour<br>Locally weighted learning<br>Learning Vector | Categorical |
| Linear | Linear regression | Continuous |
| Support vector machine | Support vector machine<br>Least squares support vector machine | Categorical<br>Continuous |
| Tree-based | Decision tree<br>Random forest<br>Gradient boosting<br>Regression tree | Categorical<br><br><br>Continuous |

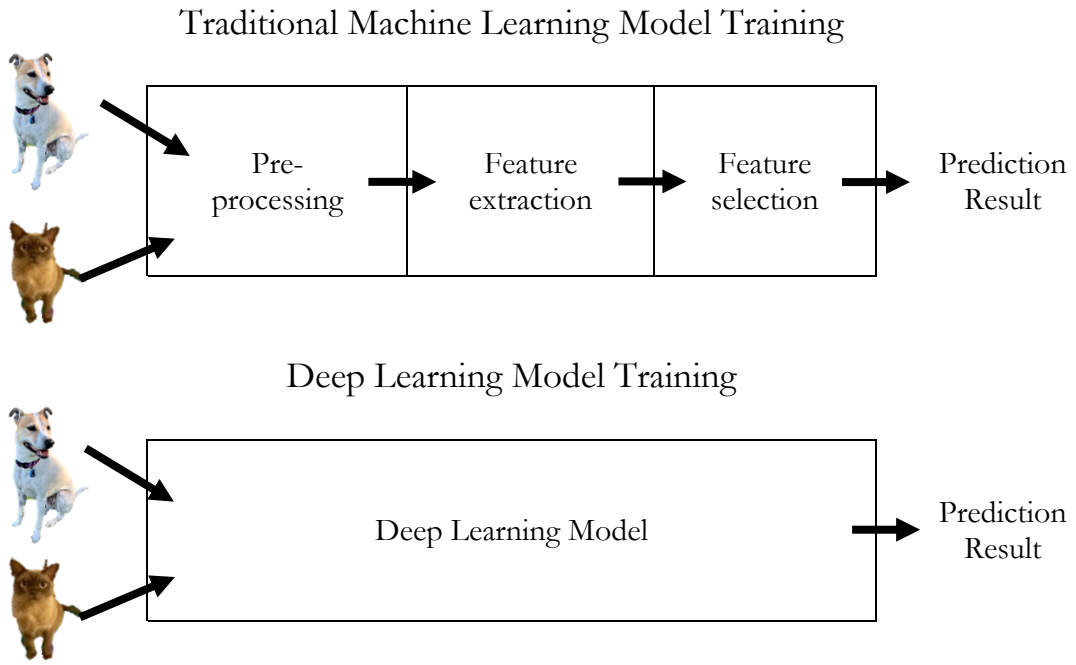
Linear regression is one of the simplest forms of traditional machine learning for continuous data, whereby a linear equation is fit to the observed independent and dependent variables. Another example is a decision tree, which can handle both continuous data for regression tasks and categorical data for classification tasks (Somvanshi et al., 2016). In a decision tree, the data is branched with each node coded for a particular decision. At each node an independent variable is input with a threshold set to make the decision that indicates the navigation path to the next branch of the tree (Somvanshi et al., 2016). Traditional machine learning models require a-priori feature selection as determined by the researcher, which is dependent on the researcher's interpretation of what features are important. Traditional approaches are also unable to handle complex interactions or relationships between inputs. A more sophisticated machine learning approach is known as deep learning, which is a subclass of machine learning that can handle more complex problems.

#### 2.5.1.1.2 Deep Learning

Deep learning is a newer form of machine learning that involves a complex architecture that is modelled on a layered structure of 'neurons' (Sarker, 2021). Like the structure of a brain, each neuron is connected to multiple other neurons, which in turn are further connected to others resulting in a deep multilayered neural network (LeCun et al., 2015; Liu & Lang, 2019; Sarker, 2021).

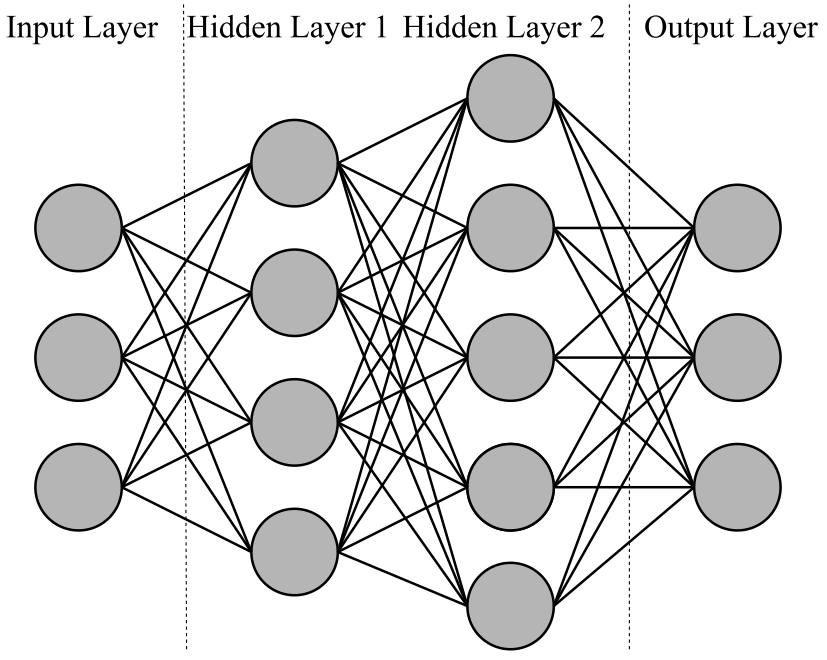
While traditional machine learning requires a human programmer to make a-priori decisions about which features are important, deep learning automates this process (**Figure 2-3**) (LeCun et al., 2015). Some researchers consider traditional approaches to have been superseded by deep learning because of the automated pre-processing, feature extraction and feature selection (**Figure 2-3**) and ability to learn complex multidimensional nonlinear relationships within the training data (LeCun et al., 2015; Jindong Wang et al., 2019; Xu et al., 2019). However, evidence of the superior performance of deep learning approaches is only just emerging for models trained IMU data (Fridriksdottir & Bonomi, 2020) and therefore requires further exploration.

**Figure 2-3. Comparison of tradition machine learning and deep learning**




In a deep learning model, there are three types of neuron layers: the input, hidden and output layers. A deep neural network is 'deep' because of multiple hidden layers, where the more hidden layers result in a deeper network. **Figure 2-4** depicts a simple neural network that is two layers deep with each grey circle represents a single neuron which is linked (or 'synapsed') to every other neuron before and after.

**Figure 2-4. Example of a deep neural network with two hidden layers**



The input layer contains the collected pre-processed data which is then passed into one or more hidden layers (Sarker, 2021). A hidden layer has multiple neurons that contain reprocessed data. Reprocessed data in each hidden neuron is a weighted sum of its inputs (Sarker, 2021). That hidden neuron then passes the new value into the next layer, with its new weighting and bias into an activation function (Sarker, 2021). The bias controls when the activation function is initiated. Once the activation function is initiated it prepares an output to further propagate the restructured data to the next layer (Sarker, 2021). Last of all, the output layer provides the results. Examples of deep learning architectures are summarised in **Table 2-5** (Liu & Lang, 2019; Sarker, 2021).

**Table 2-5. Types of supervised deep neural networks.**

| Type of Architecture | Variations | Type of data |
|---|---|---|
| Recurrent neural network | LSTM<br>Bi-LSTM<br>GRU | Regression |
| Convolutional neural network | CNN | Categorical |
| Feed-forward neural network/Artificial neural network | Multilayer perceptron | Regression |

LSTM = long-short term memory, Bi-LSTM = bidirectional-LSTM, GRU = gated recurrent units, CNN = convolutional neural network

### 2.5.2 Machine Learning in Healthcare

With the rapidly increasing digitisation of healthcare, there is a growing interest in machine learning applications for handling the large complex datasets across health disciplines such as cancer, cardiovascular disease and neurology (Jiang et al., 2017). The majority of machine learning research in healthcare is focused on diagnostic imaging, genetic profiling/diagnosis and electrodiagnosis (Jiang et al., 2017).

In healthcare, unsupervised machine learning can be used in an exploratory manner to find naturally occurring patterns in the data (Deo, 2015; Leslie et al., 2018; Sidey-Gibbons & Sidey-Gibbons, 2019). For example, it has been demonstrated that unsupervised machine learning approaches can be used for large complex multifacility datasets to improve health system performance (Leslie et al., 2018) through to monitoring outcomes at a personalised, individual level (Kobsar & Ferber, 2018). However, for the purposes of predicting outcomes, supervised machine learning

approaches have been used for broad range of healthcare applications, including diagnostics (Caballé-Cervigón et al., 2020), prognostics (Senders et al., 2018) and clinical decision making (Hassan et al., 2021).

### 2.5.3 Assessing Performance of Machine Learning Models

#### 2.5.3.1 Validation Approaches

Understanding validation approaches is important because when comparing different studies, differences between validation approaches can hinder meaningful comparison. A common validation method is k-fold validation. For example, in a 10-fold cross-validation (**Figure 2-5**), where k = 10. A k-fold cross-validation pools all the data across participants, and *randomly* partitions trials into k (10) groups (folds). From this, nine folds are used for training and one for testing. This process is repeated 10 times and the results averaged across the 10 folds.

**Figure 2-5. 10-fold cross-validation**

| Validation iteration | Training *folds* | | | | | | | | | Test **fold** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1st | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 2nd | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 3rd | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| … | | | | | | | | | | |
| 10th | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Another common approach is known as leave-one-out cross-validation (**Figure 2-6**). This again pools all participant data and randomly partitions the data across participants. But in this approach the model is trained on all the data except for a single sample on which the model is tested. Randomly partitioned k-fold validation (including both 10-fold and leave-one-out cross-validation) increase the likelihood a participant's data is part of both the training and test data. Because samples from one participant's data is potentially in both the training and test phase those approaches are known to result in higher accuracy compared to leave-one-*subject*-out cross-validation that fully takes into account between-participant variability (see below) (Gholamiangonabadi et al., 2020; Saeb et al., 2017).

**Figure 2-6. Leave-one-out cross-validation**

| Validation iteration | Training *samples* | | | | | | | | | Test **sample** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1st | 1 | 2 | 3 | 4 | 5 | 6 | … | 98 | 99 | 100 |
| 2nd | 1 | 2 | 3 | 4 | 5 | 6 | … | 98 | 99 | 100 |
| 3rd | 1 | 2 | 3 | 4 | 5 | 6 | … | 98 | 99 | 100 |
| | | | | | … | | | | | |
| 100th | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Similarly, percentage-based validation methods exist (e.g. 70% training, 30% test (**Figure 2-7**) and they also do not fully account for between-participant variability as the model is tested on, for example, 30% of the entire sample across all participants.

**Figure 2-7. 70:30 cross-validation**

| Validation iteration | Training data | Test data |
|---|---|---|
| 1 | 70% | 30% |

What is meant by 'do not account for between-participant variability' is that k--fold cross-validation and percentage split validation increases the likelihood that an individual's data is included in both the training and validation dataset (Saeb et al., 2017). That provides imprecise information about the expected accuracy for a single person who was not in the training sample, such as new patient in a clinic. In contrast, an approach where between-participant variability is maximally accounted for is known as leave-one-*subject*-out cross-validation (**Figure 2-8**).

Leave-one-*subject*-out cross-validation is similar to k-fold validation, but instead k = the number of participants (n) where, rather than randomly partitioning the data across participants, *each participant's data* is partitioned. The model is trained on data from all participants except for that of the test participant and the results of the model are averaged across the number of participants.

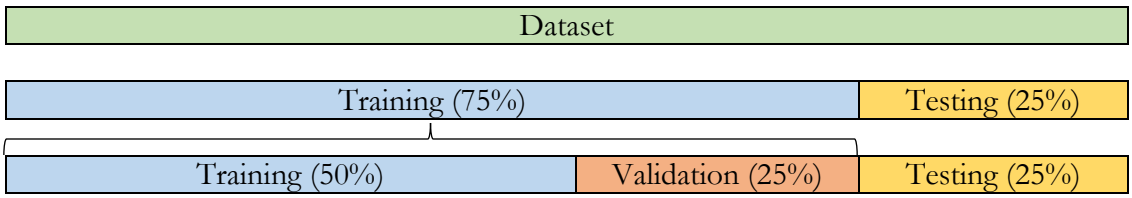
**Figure 2-8. Leave-one-subject-out cross-validation**

| Validation iteration | Training *participants* (**n - 1**) | | | | | | | | | Test **1** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1st | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 2nd | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 3rd | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| | | | | | … | | | | | |
| 10th | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

*Sidenote*

Some researchers split data into three sub-sets – training, validation, and test (**Figure 2-9**). In these instances, there is the additional validation step which subdivides the training data, providing an opportunity to refine model parameters to optimise the model prior to testing. For example, Fridriksdottir and Bonomi (2020) used 50% training data, 25% validation, and 25% test to validate a human activity recognition model.

**Figure 2-9. Validation sub-set of training data**




Gholamiangonabadi et al. (2020) reported the results for human activity recognition model reporting both 10-fold cross-validation (99.8% accuracy) and leave-one-*subject*-out cross-validation (85.1% accuracy). Those results demonstrate a substantial ~15% difference between the two validation methods. For a clinician, it is difficult to interpret how accurate a human activity recognition system is for each new patient, when accuracy is derived from a 10-fold cross-validation rather than a leave-one-subject-out cross-validation. This is because the results of a 10-fold cross-validation are confounded by training data from a test participant potentially being in the training data. A leave-one-subject-out cross-validation never uses data from a test participant in the training data. Therefore, when a machine learning model is used for a new patient in a clinical practice, a clinician would know the average expected accuracy for that patient, presuming the patient has the same characteristics as the participants on which the model was built. And that is quite useful for clinicians.

#### 2.5.3.2 Statistical Reporting

Across different fields of research, language for statistical testing differs. In medical research, accuracy is commonly reported for diagnostic tests whereby the results of a test can be cross tabulated into four categories: true positive, false positive, true negative and false negative. The cross-tabulation (otherwise known as a confusion matrix), provides a summary of the frequency of possible outcomes. From the frequencies, various statistics can be derived to describe accuracy (**Table 2-6**).

The same confusion matrix is used for machine learning classifiers, although different statistics are used.

**Table 2-6. Confusion matrix.**

| | | Actual | | |
|---|---|---|---|---|
| | | Positive | Negative | |
| Predicted | Positive | True Positive | False Positive | Precision |
| | Negative | False Negative | True Negative | Negative Predictive Value |
| | | Sensitivity/Recall | Specificity | Overall Accuracy |

Accuracy in the context of diagnostics is commonly reported as (Alberg et al., 2004):

**Sensitivity**: the number of positive tests out of all those who have the disease.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

**Specificity**: the number of negative tests out of all those who don't have the disease.

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive}$$

Sensitivity and specificity are most commonly used for diagnostics because 'overall accuracy' (described below) is dependent on the known prevalence of a condition which affects precision (Alberg et al., 2004). However, in machine learning validation, the exact number of positive samples is always known, and therefore 'overall accuracy' is appropriate for use. Another difference regarding machine learning, is that the context for interpretation is about predicting correct positives, whereas in clinical diagnostics correctly identifying true negatives is often equally important, and therefore instead of specificity, precision is used. The one metric that is similar between diagnostics and machine learning is recall/sensitivity which describes the correctly identified positives out of all the positives that exist.

Accuracy in the context of machine learning classification can be reported as:

**Overall Accuracy**: the number of correct predictions out of the total number of predictions.

$$Overall\ Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative}$$

**Precision**: the number of correct positive predictions out of the total number of positive predictions.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

**Recall**: the number of correct positive predictions out of the actual positive cases.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

### 2.5.4 Supervised Machine Learning in Knee Osteoarthritis

Since 2015, supervised machine learning approaches have attracted significant increase in attention from knee osteoarthritis researchers (Jamshidi et al., 2019; Kokkotis et al., 2020). A recent literature review reported the two most common types of training data used for machine learning models were medical imaging (e.g. x-ray or magnetic resonance imaging) and biomechanical data (e.g. from IMUs) (Kokkotis et al., 2020). At times these were combined with patient-reported outcome measures (e.g. KOOS), demographic data, health status, genetic data, biochemical markers and food intake (Jamshidi et al., 2019; Kokkotis et al., 2020). In the field of knee osteoarthritis, there has been significant interest in using supervised machine learning approaches for diagnostics and prognostics with little investigation into the development of tools that can assist with clinical decision making (Jamshidi et al., 2019; Kokkotis et al., 2020). Amongst the 75 studies in the review by Kokkotis et al. (2020); 13 studies used prediction or regression techniques that usually used medical imaging to predict structural progression; 43 studies used classification techniques using medical imaging or biomechanical data for the purposes of diagnostics; 15 studies used machine learning techniques for segmenting medical imaging; and four studies used biomechanical data for monitoring rehabilitation. The majority of the machine learning-osteoarthritis research agenda is clearly weighted towards

diagnostics, rather than the development of tools to facilitate clinical decision-making post-diagnosis. It is clear that there is a gap in the literature for using machine learning approaches to handle IMU data in a way that would provide clinicians information from which to base decisions about management of their patients.

#### 2.5.4.1 Machine Learning for IMU Data in People with Knee Osteoarthritis

For a person with knee osteoarthritis, IMU-machine learning approaches have the potential to provide a clinician or researcher with highly individualised, clinically relevant information about movement related outcomes. Despite this, there has yet to be significant research interest in machine learning approaches to facilitate biomechanical analysis and activity monitoring for people with knee osteoarthritis in free-living environments.

Machine learning models that use IMU data are well placed to address the limitations with current methods of monitoring outcomes in clinical practice (see section 2.3) as they have the potential to provide objective data about how a person moves and engages with activities in free-living environments. Machine learning approaches also avoid the limitations associated with fusion algorithms outlined in section 2.4.2. by being able to use raw accelerometer and/or gyroscope data while ignoring the magnetometer data.

There are a limited number of studies that have investigated the development of machine learning IMU monitoring systems for people who have knee osteoarthritis. Kobsar et al. (2017) reported ~80% accuracy classifying high-responders, low-responders, and non-responders on the KOOS following a hip strengthening intervention from gait data (four accelerometers) using a traditional machine learning approach (principal component analysis). Traditional machine learning approaches have also been used to classify common rehabilitation exercises prescribed for people who have knee osteoarthritis (Chen et al., 2015; Huang et al., 2017). These studies use two IMUs placed on the lower limb and up to three IMUs also placed elsewhere on the body of healthy participants to classify three different exercises and reported the accuracy of these models as ranging from 62% to 98%. While these machine learning IMU-based systems would be helpful for monitoring correct performance of an exercise, or compliance with an exercise programme in free-living environments, they

are unable to assist in monitoring other clinically important aspects of physical function such as biomechanics and monitoring of daily activities (e.g. walking, negotiating stairs, and transitioning to and from a chair) (see section 2.3.2).

So far, there are few studies investigating development of a machine learning IMU human activity recognition system for activities commonly reported as difficult in people with knee osteoarthritis (Emmerzaal et al., 2020; Lipperts et al., 2017; Verlaan et al., 2015) (see section 2.5.5) – although those studies did not validate models on the intended population. There are also only a handful of studies that have developed machine learning models to predict biomechanical outcomes for functional activities (He et al., 2019; Renani et al., 2021; Renani et al., 2020; Wang et al., 2020) (see section 2.5.6).

There is a clear, untapped potential for future development of machine learning models trained on IMU data for human activity recognition and prediction of movement parameters. The combination of IMU technology with data handling using machine learning approaches has the potential to aid clinical decision making for clinicians managing people who have knee osteoarthritis. The following sections will explore machine learning approaches to human activity recognition and biomechanical prediction.

### 2.5.5 Machine Learning for Human Activity Recognition

There are currently only two methods of assessing physical function outside of a clinical or laboratory environment: using patient-reported outcome measures (see section 2.3.1), and using accelerometers to monitor energy expenditure or body position (see section 2.3.2.3). But neither of these approaches provides direct information about the actual performance of activities that are most commonly associated with activity limitation due to knee osteoarthritis, which include walking, negotiating stairs and transitioning to or from a chair (Dobson et al., 2013). The goal of machine learning based human activity recognition is to use movement-based data to recognise and classify activities performed by humans in free-living environments (Kim et al., 2010).

A human activity recognition system designed to classify clinically important activities would have potential to assist in clinical decision making by monitoring performance of these activities when a patient is unobserved, outside of the clinical or laboratory environment. Because of the limitations associated with monitoring outcomes in clinical practice (see section 2.3), human activity recognition systems provide a promising new tool that has the potential to support clinical decision making.

Monitoring systems have the potential to provide a clinician information about the frequency their patient has performed, and/or the time spent engaging with, an activity in free-living environments (Verlaan et al., 2015). Data from a human activity recognition system could be used by a clinician, for example, to reduce avoidance of activities by providing motivating feedback from monitoring data about active time versus sedentary time, or engagement with feared or painful activities when at home or out in the community. In addition, a second use for human activity recognition systems could be to segment large data streams into datasets of specific activities for subsequent biomechanical analysis.

Human activity recognition models are developed as supervised machine learning classifiers, like the earlier example about classifying dogs and cats (see section 2.5.1). For human activity recognition, the data are categorised and labelled before being input to train the machine learning model. A range of technologies exist that output data which could be input into a human activity recognition system (Qi et al., 2018). Devices are categorised as 'on-body' (wearable) sensors (e.g. IMUs or global positioning systems) or 'on-object' sensors (e.g. radiofrequency identification tags or infrared location devices). On-body sensors are not limited by location in that they can be worn in any environment (i.e. indoors, outdoors, home, work, recreation). On-body technology such as instrumented shoe insoles (Ngueleu et al., 2019), body cameras and IMUs (Qi et al., 2018) are available for the purposes of human activity recognition. Instrumented insoles have limitations because they have the additional burden of needing to be changed between the shoes worn and they do not fit all shoes. While body-worn cameras can provide direct observation of an activity, there are privacy concerns that influence user acceptance because they record the user's

environment. In contrast, IMUs do not have these limitations, although wearability limitations exist as some people may find them uncomfortable or cumbersome. Despite this, IMUs are reducing in size and weight and can be considered a practical option for collection of data in free-living environments. However, IMU data collected in free-living environments are typically unlabelled, posing a data handling challenge. Human activity recognition can be the first part of a data handling pipeline that automatically segments data into manageable samples, providing clinically relevant information.

The clinical application for human activity recognition for people with knee osteoarthritis falls into two categories: (a) classifying rehabilitation exercise, and (b) classifying functional activities. The literature review by Kokkotis et al. (2020) identified two studies that developed traditional machine learning human activity recognition models (support vector machines) to classify rehabilitation exercises commonly prescribed to people who have knee osteoarthritis (Chen et al., 2015; Huang et al., 2017). The exercises selected in these studies included; sitting knee extension, supine straight leg raise, and standing hip abduction (Huang et al., 2017); and inner range knee extension, supine straight leg raise and a mini-squat (Chen et al., 2015). The classification accuracy for these studies ranged from 99.3% (SD 1.16%) (Chen et al., 2015) to 100% (Huang et al., 2017) using a 10-fold cross-validation. These studies have limited clinical utility for three reasons. Firstly, they recruited young, healthy participants rather than people who have knee osteoarthritis. Secondly, while having a system that can classify exercises might be helpful for monitoring exercise training volume during rehabilitation, it does not provide information about a person's functional status. Thirdly, because those studies use a 10-fold cross validation that does not provide a clinician information about the expected average accuracy for each individual patient (see section 2.5.3.1).

There are a substantial number of studies that have investigated the development of machine learning models to classify functional activities from inertial measurement unit data collected from healthy people (Arif & Kattan, 2015; Ascioglu & Senol, 2020; Bulling et al., 2014; Cust et al., 2019; Fridriksdottir & Bonomi, 2020; Martinez-Hernandez & Dehghani-Sanij, 2018, 2019; O'Reilly et al., 2018; Ramanujam

et al., 2021; Song-Mi et al., 2017; Jindong Wang et al., 2019; Weygers et al., 2020; Xu et al., 2019; Xu et al., 2018). Fewer studies have investigated populations that have pathology affecting lower limb movement (Albert et al., 2012; Lonini et al., 2016; Rast & Labruyère, 2020). Human activity recognition studies differ significantly on the types of activities, the number of activities, if the activities involve both upper and lower limbs, the number of IMUs used, the placement of IMUs, and the validation approach used. For example, Arif and Kattan (2015) reported a 89% accuracy (70% training data, 30% test validation – see **Figure 2-7**) classifying 12 activities (lying, sitting, standing, walking, running, cycling, Nordic walking, ascending stairs, descending stairs, vacuuming, ironing and rope jumping), by training a neural network on data collected from nine participants wearing three sensors (wrist, chest, and ankle). In comparison Fridriksdottir and Bonomi (2020) reported a 94% accuracy (50% training data, 25% validation, and 25% test – see **Figure 2-9**) classifying six activities (lying, upright, walking, descending stairs, ascending stairs and using a wheel chair), by training a deep neural network on data from 20 participants wearing a single sensor placed on the trunk. The heterogeneity in the type and number of activities, number of IMUs, placement of IMUs, and validation approaches makes direct comparison of the accuracy between human activity recognition models challenging. Nonetheless, for a human activity recognition system to be clinically meaningful, it is optimal for training data to be trained to classify only a limited number of activities relevant to the patient's pathology. Limiting the number of activities ensures the data from the human activity recognition system is interpretable for the clinician. For example, some human activity recognition systems are designed to classify different standing activities such, as opening a refrigerator (Ascioglu & Senol, 2020) or ironing (Arif & Kattan, 2015). But this differentiation between standing activities is arguably unimportant for a clinician managing a person who has knee osteoarthritis because they are both standing activities that are largely differentiated by arm movement (which is not the focus for people with knee osteoarthritis) and time spent performing the activity.

In a scoping review, Rast and Labruyère (2020) investigated the use of IMUs for quantifying activities in people with mobility impairments. That review identified 95

studies that used various forms of human activity recognition that could be categorised into four categories: "(1) maintaining and changing a body position, (2) walking and moving, (3) moving around using a wheelchair, and (4) activities that involve the upper extremity)" (Rast & Labruyère, 2020). The categories most consistent with the OARSI recommendations for assessment of physical function in people who have knee osteoarthritis (Dobson et al., 2013) relate to category one (changing body position, e.g. transitioning to and from a chair) and category two (walking and moving, e.g. negotiating stairs and walking) because of the primary physical limitation in moving the lower limbs. Machine learning based human activity recognition for people with knee osteoarthritis should focus on those activity categories. Human activity recognition machine learning models focused on lower limb specific activities have been developed for people with knee osteoarthritis, or post knee replacement. These three studies used traditional machine learning models (decision tree) to classify clinically important activities such as walking, ascending stairs, descending stairs, and transitioning to and from a chair (Emmerzaal et al., 2020; Lipperts et al., 2017; Verlaan et al., 2015).

Using an unspecified validation approach, Lipperts et al. (2017) reported a 99.5% prediction accuracy (five errors from 992 samples) for classifying activities for healthy participants and 98.9% for people who had a (unspecified) joint replacement (four errors from 390 samples). That study recruited 16 healthy participants and 40 participants who had received a joint replacement three to 14 days prior to testing who were all asked to walk around, sit, stand, ascend, and descend stairs. Because that study did not specify the validation approach, it is not possible to interpret their model performance against similar studies. Emmerzaal et al. (2020) developed an app designed for people with knee osteoarthritis that included human activity recognition and biomechanical analysis to facilitate clinical decision making. They trained an unspecified machine learning model to classify the activities of walking, ascending stairs, descending stairs, sit-to-stand, stand-to-sit, jogging, and cycling, by use of data measured by a mobile phone accelerometer in a hip bag. The accuracy predicting these activities ranged between 65% to 97% using a leave-one-subject-out cross-validation approach (Emmerzaal et al., 2020). However, their model was trained and

tested on data from 17 healthy participants, leaving questions about the validity of the human activity recognition model for the intended population who have knee osteoarthritis. The study by Verlaan et al. (2015) included participants with end-stage knee osteoarthritis, but their aim was to compare performance of physical activities between that group and healthy controls, rather than to develop and validate a human activity recognition system, and therefore did not report validation metrics precluding meaningful interpretation of their model's performance. Overall, human activity recognition systems have been tested for clinically important activities but have not been trained nor tested on a population with knee osteoarthritis which is an important step prior to clinical implementation.

### 2.5.6 Machine Learning for Predicting Movement Parameters

IMUs also have the potential to be able to monitor movement in people with knee osteoarthritis for the purpose of biomechanical analysis. However, there are currently some limitations, and these have contributed to the slow uptake in clinical practice (see section 2.4.2). Over the past few years, there has been an increasing number of papers (below) investigating deep learning approaches for predicting biomechanical outcomes from IMU data as a method to overcome the calibration and magnetisation limitations of this technology. Machine learning based biomechanical prediction models have been developed for the hip, knee and ankle and have included both kinematic and kinetic movement parameters.

Knee joint movement parameters predicted in machine learning studies include spatiotemporal (Renani et al., 2020) and angular kinematics (Hernandez et al., 2021; Rapp et al., 2021; Wouda et al., 2018), as well as joint moments and forces (He et al., 2019; Mundt et al., 2021; Mundt, Koeppe, David, Witter, et al., 2020; Mundt, Thomsen, et al., 2020; Stetter et al., 2020; Stetter et al., 2019; Wang et al., 2020). Most studies about the development of machine learning prediction models for knee joint movement parameters usually include young and healthy participants. In these studies that recruit healthy participants, a variety of deep learning approaches have been used including artificial neural networks/feed-forward neural networks (Mundt, Thomsen, et al., 2020; Wouda et al., 2018), long-short term memory (LSTM)/recurrent neural networks (Mundt et al., 2021; Mundt, Thomsen, et al., 2020; Rapp et al., 2021),

multilayer perceptron networks (Mundt et al., 2021), convolutional neural networks (CNN) (Mundt et al., 2021), and combined CNN-LSTM (Hernandez et al., 2021). These studies developed biomechanical prediction models only for the activity of walking or sports-specific lower limb movements.

Machine learning biomechanical prediction models are usually compared to gold-standard optoelectronic motion analysis systems for kinematic parameters, with the addition of force plate data for kinetic parameters. **Table 2-7** outlines the prediction error for various deep learning prediction models for knee kinematic and kinetic parameters for walking.

**Table 2-7. Deep learning model prediction error for walking.**

| Movement parameter | r | RMSE | Study |
|---|---|---|---|
| Sagittal plane angular kinematics | | | |
| Flexion/ extension | 0.94 – 0.99 | 0.97 – 12.1° | (Hernandez et al., 2021; Mundt et al., 2021; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Wouda et al., 2018) |
| Knee moments | | | |
| Adduction | 0.71 – 0.98 | 10.5 – 22.3%* | (He et al., 2019; Mundt, Thomsen, et al., 2020; Stetter et al., 2020) |
| Flexion | 0.72 | 18.4%* | (Stetter et al., 2020) |
| Knee joint forces | | | |
| Medial | 0.6 | 27.7%* | (Stetter et al., 2020) |
| Compression | 0.87 | 14.2%* | |

*Normalised RMSE = RMSE divided by the peak to peak amplitude (Ren et al., 2008), r = Pearson's correlation coefficient.

There has been little in the way of machine learning research to predict movement parameters using IMU training data collected from people who have knee osteoarthritis despite the potential clinical significance of movement patterns in this population (see sections 2.1.3.1, 2.1.4, 2.2.1, 2.3.2.2). A recent review by Kobsar et al. (2020) investigated the use of wearable sensors for gait analysis in people who have knee osteoarthritis. From 72 studies, they identified only two studies using machine learning for predicting knee joint moments (He et al., 2019; Wang et al., 2020), and no studies for the prediction of kinematics (Kobsar et al., 2020). Only two additional studies since that review have investigated spatiotemporal (Renani et al., 2020), and

kinematic gait parameters (Renani et al., 2021) using data from participants with knee osteoarthritis.

The two studies identified in the review by Kobsar et al. (2020) used training data collected from people with knee osteoarthritis wearing IMUs placed on both ankles to train an artificial neural network to predict knee adduction moment during the stance phase of walking. Wang et al. (2020) demonstrated the mean absolute error for their ANN to be 0.004 Nm/kg*cm and $R^2 = 0.96$, while He et al. (2019) reported an RMSE of 0.36 Nm/kg*m (SD = 0.11) and $r = 0.91$. Because of the different methods of reporting results, it is not possible to directly compare these two studies, nor make comparisons between these studies that recruited people with knee osteoarthritis and other studies that only recruited healthy people.

Renani et al. (2021) trained a bidirectional LSTM (BiLSTM) (deep learning) model to predict knee and hip angular kinematics (sagittal, frontal, and transverse planes) and spatiotemporal parameters (segment angular velocity and acceleration) for walking. They used training data collected from four IMUs placed over the pelvis, thigh, shank, and foot of participants, 13 of whom had knee osteoarthritis and 17 had a previous total knee arthroplasty. Averaged across the three predicted planes, they reported RMSE (SD) for hip 4.5° (1.6°) and knee 3.3° (0.2°) with strong correlations for the hip $r = 0.82$ and knee $r = 0.83$. The most accurate prediction was for sagittal plane knee movement (RMSE 2.9° (1.1°), $r = 0.99$).

The prediction error for the deep learning kinematic prediction models that use IMU data collected from people with knee osteoarthritis for training and testing (He et al., 2019; Renani et al., 2021; Wang et al., 2020) are similar to that reported in other studies using data collected from healthy participants (Hernandez et al., 2021; Mundt et al., 2021; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Wouda et al., 2018). This demonstrates the feasibility for training and testing machine learning models to predict knee joint biomechanics for walking in this clinical population. Although the error may be similar, this does not detract from the need to derive models using data from people with activity limitation or pain, as their movement patterns can be different, potentially impacting model performance (as detailed in the next section).

No studies that have trained machine learning models on data from participants with knee osteoarthritis have demonstrated performance for predicting kinematic or kinetic parameters for multiple clinically relevant activities (i.e. walking, negotiating stairs and transitioning to and from a chair). Because biomechanical prediction models have only been developed for walking for people who have knee osteoarthritis, it is unknown if similar models can be built across a range of clinically important activities with acceptable accuracy. The performance of a model that is trained on data from a range of activities including walking, negotiating stairs, and transitioning to and from a chair, remains unclear.

### 2.5.7 Population-Specific Models

While there is a growing body of research of machine learning prediction studies using IMU training data from healthy participants, the models developed within these studies may not be fit for purpose in clinical populations, such as people with knee osteoarthritis. Very few studies have developed machine learning models trained and tested on data from clinical populations with activity limitation or pain (Rast & Labruyère, 2020).

For some studies that use data from healthy participants, the intended end use is for populations with specific pathology such as knee osteoarthritis. This is potentially problematic, as people with knee osteoarthritis move differently to those without knee osteoarthritis (see section 2.1.4). Because machine learning predictions rely on consistent patterns in the IMU data, using models built on training data from healthy participants can affect the accuracy of the predictions when the model is tested on populations with pathology that affects their movement.

There is some early work that has revealed machine learning prediction models using IMU data are affected by the population from which the training data was collected. For example, in human activity recognition studies, studies report 11% to 26% reduction in test accuracy for models trained on data from healthy people that are subsequently tested on people with movement impairments such as Parkinson's disease (Albert et al., 2012) or who are wearing orthoses (Lonini et al., 2016).

As described in the previous section 2.5.5, Emmerzaal et al. (2020) developed a human activity recognition system for people with knee osteoarthritis as one component of an IMU-app. However, they trained and tested their model on data from healthy participants, but not people with knee osteoarthritis. After developing the human activity recognition model, they applied the model in a subsequent test for people who have knee osteoarthritis to explore the usability of their product. Their study did not report the test results on data from people with knee osteoarthritis but instead, the authors described that during usability testing in people with knee osteoarthritis, their system was inaccurate, limiting their systems utility (Emmerzaal et al., 2020). These findings may suggest that the difference in movement patterns between healthy participants and those with knee osteoarthritis affected the results. Future studies should therefore consider training and testing human activity recognition models on data collected from people with knee osteoarthritis.

It is also not known if machine learning biomechanical prediction models can be used broadly across different populations. A recent study reported on the development of a deep learning model for spatiotemporal kinematic predictions, using training data from 14 participants with knee osteoarthritis and 15 participants who had a total knee arthroplasty (Renani et al., 2020). They reported a 4% higher prediction error for people who have knee osteoarthritis compared to those who had a total knee arthroplasty. They concluded that the difference in prediction error was because of greater variability in gait parameters for participants with knee osteoarthritis compared to those who had a knee arthroplasty.

Considering there are significant differences in kinematics and kinetics between people with and without knee osteoarthritis (see section 2.1.4), machine learning models for human activity recognition and biomechanical prediction may need to be trained on population specific data to minimise prediction error.

### 2.5.8 Key Points

*What is known and not known about the use of artificial intelligence in studying people with knee osteoarthritis?*

- Machine learning is a branch of artificial intelligence where through experience a model can be trained to predict an outcome.
- Deep learning is a branch of machine learning that does not require a-priori feature selection by the researcher and is capable of handling non-linear relationships.
- In people with knee osteoarthritis, supervised machine learning has been used for predicting diagnosis and prognosis with minimal focus on models to facilitate clinical decision making.
- Machine learning can be used as an alternative to fusion algorithms for IMU data to overcome limitations with electromagnetic interference and calibration requirements.
- There are no studies using machine learning human activity recognition in people with knee osteoarthritis for activities associated with activity limitation.
- There are a few studies using machine learning for biomechanical prediction in people with knee osteoarthritis but only for walking.
- Population-specific models may be required.

## 2.6 Summary of Literature

Knee osteoarthritis is a growing problem globally and accounts for the largest proportion of the total osteoarthritis burden and its associated disability. That disability is related to symptoms such as persistent pain and stiffness, psychological factors and physical factors that are associated with movement impairment and activity limitation. The most common activity limitations include walking, negotiating stairs, and transitioning to and from a chair, impacting a person's ability to fully participate in society. People with knee osteoarthritis display altered movement patterns when performing clinically relevant activities such as walking, negotiating stairs, and transitioning to and from a chair. Movement parameters most frequently observed to differ from healthy controls include increased knee adduction moment, reduced sagittal plane range of movement, and increased levels of muscular activity around the knee. Abnormal knee joint loading, such as knee adduction and flexion moments, are risk factors for structural progression of knee osteoarthritis.

A variety of non-invasive interventions (e.g. braces, orthotics, exercises) have been developed that aim to reverse the change in movement patterns in people with knee osteoarthritis in an effort to prevent structural progression, and improve pain and resultant activity limitation. Exercise is a core guideline-based intervention for people with knee osteoarthritis. Multiple high-quality studies demonstrate that exercise improves activity limitation and pain in people with knee osteoarthritis. Some exercise approaches have demonstrated the potential to influence movement parameters, although, there is conflicting information about whether exercise interventions can change movement patterns.

There is one systematic review investigating the relationship between changes in movement and changes in symptoms, which included randomised controlled trials and only investigated knee adduction moment during walking. They concluded that there is no relationship between changes in knee adduction moment and clinical outcomes assessed with patient-reported outcome measures. However, other activities, movement parameters and study designs have yet to be systematically

investigated to explore the relationship between changes in movement and changes in activity limitation and pain after exercise interventions.

There is heterogeneity across the population with knee osteoarthritis. Because movement patterns are variable across the population, group-based intervention and assessment may disguise the relevance of targeting a change in individualised movement parameters through exercise-based interventions. There are two recent but small studies that have demonstrated strong relationships between changes in individual person-level movement parameters, assessed using IMUs, and changes in clinical outcomes.

IMUs are one tool that when combined with machine learning approaches have the potential to be used for monitoring the physical activities of patients with knee osteoarthritis in free-living environments, which could be used to help guide clinical decision making. However, IMUs systems are prone to error in free-living environments because they rely on magnetometers that can be affected by electromagnetic interference from computers, mobile phones, and metallic structures. Machine learning approaches to processing IMU data using only the raw accelerometer and gyroscope data have been successfully used in previous studies for two specific and clinically relevant purposes: human activity recognition and biomechanical prediction.

Machine learning is also one method that can handle the large amounts of data produced by IMUs worn in free-living environments. But as the recorded data is unlabelled, it typically provides outputs that do not describe the type of activity that was being performed, preventing meaningful analysis for clinicians. Human activity recognition is a machine learning approach that could be used for monitoring engagement with activities in free-living environments, whist also segmenting the data for subsequent biomechanical analysis. Several studies have used IMU data for training biomechanical prediction models in people with knee osteoarthritis. But currently, no studies have demonstrated the prediction error for machine learning models for any activity besides walking and the majority use training data collected from healthy participants.

Most IMU-machine learning studies have used IMU data collected from healthy participants, limiting generalisability for clinical populations such as those with knee osteoarthritis. While previously published human activity recognition and biomechanical prediction machine learning models based on training data from healthy participants could be used for people with knee osteoarthritis, the significant differences in movement patterns between people with and without the condition may limit their validity and reliability. So far there are no human activity recognition machine learning studies, and only a handful of studies describing machine learning based biomechanical prediction models for kinematics and kinetics, that have used IMU training data collected from people with knee osteoarthritis.

It is unclear if it is feasible to develop a machine learning IMU-based human activity recognition and biomechanical prediction system for multiple clinically important activities, based on IMU training data collected from people with knee osteoarthritis. Such a system could be used in free-living environments to monitor individualised changes in engagement with clinically important activities and to monitor changes in movement patterns that result from exercise, or other interventions for people with knee osteoarthritis.

## 2.7 Aims of Thesis

1. Systematically review cohort studies and randomised controlled trials to investigate how changes in knee joint movement parameters during functional activities relate to changes in activity limitation or pain after exercise intervention in people with knee osteoarthritis.
2. Investigate how wearable sensor technology could be used to monitor activity avoidance and altered movement patterns in people with knee osteoarthritis:
    a. Develop an IMU-based, human activity recognition system that can classify clinically relevant activities and phases of activities (walking, negotiating stairs, and transitioning to and from a chair) for people with knee osteoarthritis;
    b. Develop machine learning prediction models for knee joint sagittal plane angular kinematics for multiple clinically important activities; and
    c. Develop machine learning prediction models for knee joint moments and forces for the stance phase of walking.

# Chapter 3

# Study 1: Systematic Review

## The Relationship Between Changes in Movement and Activity Limitation or Pain in People with Knee Osteoarthritis: A Systematic Review

Prior to this body of research, there was one previous systematic review summarising the results of studies that investigated the relationship between changes in movement patterns and changes in the clinical outcomes of activity limitation or pain after exercise interventions (Ferreira et al., 2015). That review included studies that (a) investigated knee adduction moment, (b) limited study designs to randomised controlled trials, and (c) did not investigate the relationship using within-group change or correlation analyses. There are a broad range of other movement parameters that have been investigated to change post exercise. Also, both randomised controlled trials and cohort study designs are appropriate to investigate the relationship between change in outcomes using the either the co-occurrence of group-level change data or the correlation between changes using individual person-level data.

Therefore, as there were no systematic reviews that had investigated a relationship between movement patterns and clinical outcomes across a broad range of kinematic, kinetic and muscle activity parameters, questions remained about the relevance of changing movement patterns in clinical practice.

The aim of this systematic review was to summarise the evidence about the relationship between change in knee joint angular kinematics, moments and muscle activity and change in activity limitation or pain following exercise intervention in people with knee osteoarthritis.

**This chapter was published in the *Journal of Orthopaedic & Sports Physical Therapy*.**

**Tan, J.-S.**, Tikoft, E., O'Sullivan, P., Smith, A., Campbell, A., Caneiro, J. P., & Kent, P. (2021). The Relationship Between Changes in Movement and Activity Limitation or Pain in People With Knee Osteoarthritis: A Systematic Review. *Journal of Orthopaedic & Sports Physical Therapy*, *51*(10), 492-502. https://doi.org/10.2519/jospt.2021.10418

Reproduced with permission. Copyright © Journal of Orthopaedic & Sports Physical Therapy ®, Inc.

This chapter contains the submitted manuscript and includes stylistic changes to match this thesis that differ from the published version. Such changes do not affect the interpretation of results.

## 3.1 Abstract

**Objective:** To report whether changes in knee joint movement parameters recorded during functional activities relate to change in activity limitation or pain after an exercise intervention in people with knee osteoarthritis.

**Design:** Aetiology systematic review.

**Literature Search:** Four databases (MEDLINE, Embase, CINAHL, and AMED) were searched up to January 22, 2021.

**Study Selection Criteria:** Randomised controlled trials or cohort studies of exercise interventions for people with knee osteoarthritis that assessed change in knee joint movement parameters (moments, kinematics, or muscle activity) and clinical outcomes (activity limitation or pain).

**Data Synthesis:** A descriptive synthesis of functional activities, movement parameters, and clinical outcomes.

**Results:** From 3182 articles, 22 studies met the inclusion criteria, and almost all were of low quality. Gait was the only investigated functional activity. After exercise, gait parameters changed 26% of the time, and clinical outcomes improved 90% of the time. A relationship between group-level changes in gait parameters and clinical outcomes occurred 24.5% of the time. Two studies directly investigated an individual-level relationship, reporting only one significant association out of eight correlations tested.

**Conclusion:** Most studies reported no change in gait-related movement parameters despite improvement in clinical outcomes, challenging the belief that changing movement parameters is always clinically important in people with knee osteoarthritis.

## 3.2 Introduction

When walking (van Tunen et al., 2018), transitioning from sit-to-stand (Sonoo et al., 2019), or negotiating stairs (Iijima et al., 2018), people with knee osteoarthritis have altered joint loading (e.g. knee adduction moment) (Heiden et al., 2009; Iijima et al., 2018; Rutherford et al., 2017; Sparkes et al., 2019; van Tunen et al., 2018) less sagittal plane range of movement (Baliunas et al., 2002; Bouchouras et al., 2015; Hinman et al., 2002; McCarthy et al., 2013) and altered muscle activity compared to people without osteoarthritis (Bouchouras et al., 2015; Heiden et al., 2009; Hinman et al., 2002). Some altered movement patterns have been associated with structural progression of osteoarthritis (Bennell et al., 2011; Chehab et al., 2014; Thorp et al., 2006) and pain intensity (Bensalma et al., 2019; Hall et al., 2017; O'Connell et al., 2016). Gait alterations that include increased hamstrings-quadriceps co-contraction are associated with knee joint effusion (Rutherford et al., 2012), a factor that is strongly and independently related to weight-bearing knee pain in people with knee osteoarthritis (Lo et al., 2009).

It is unclear whether altered movement patterns are the cause of activity limitation or pain. Yet, rehabilitation strategies focus on changing movement patterns to reduce synovitis and structural progression, in order to subsequently improve activity limitation and pain (Ageberg & Roos, 2015; Al-Khlaifat et al., 2016; Lehman, 2018; Radzimski et al., 2012; Richards et al., 2017).

There is preliminary evidence from randomised controlled trials (Cheung et al., 2018; Hunt et al., 2018) and uncontrolled studies (Hunt & Takacs, 2014; Shull, Silder, et al., 2013; Thorp et al., 2010) that exercise may change knee movement patterns (Al-Khlaifat et al., 2016; Preece et al., 2016). There is strong evidence from multiple systematic reviews that exercise improves the clinical outcomes of activity limitation and pain (Bannuru et al., 2019; Fransen et al., 2015). However, it is unclear whether there is a relationship between changes in specific knee joint movement parameters and changes in clinical outcomes.

The relationship between change in knee adduction moment and change in clinical outcomes after an exercise intervention has been systematically reviewed

before: changes in activity limitation or pain were not associated with changes in knee adduction moment (Ferreira et al., 2015). However, the review had two important limitations: (a) it included knee adduction moment only (no other movement parameters), and (b) it included randomised controlled trials only, despite other designs (such as cohort studies) also being capable of providing insight into the relationship between a change in movement parameter and a change in clinical outcome. A more comprehensive approach is to include both randomised controlled trial and cohort study designs, using within-group statistical approaches across a comprehensive range of knee joint movement parameters.

Therefore, we aimed to assess whether changes in knee joint movement parameters during functional activities were associated with changes in activity limitation or pain after an exercise intervention in people with knee osteoarthritis.

## 3.3 Methods

We followed the reporting standards of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher et al., 2009) and prospectively registered the review protocol with PROSPERO (registration number CRD42020160164).

### 3.3.1 Data Sources

Electronic databases (AMED, CINAHL, Embase, and MEDLINE) were searched from inception to January 22, 2021. The search strategy, using Medical Subject Headings (MeSH) (National Library of Medicine, 2019) and key words, was developed in collaboration with a faculty librarian (**Appendix 3-1**).

### 3.3.2 Study Selection

We included randomised controlled trials and cohort studies that (a) studied people with clinically (e.g. American College of Rheumatology classification criteria (Altman et al., 1986)) or radiologically diagnosed symptomatic knee osteoarthritis who participated in an exercise intervention, and (b) reported outcomes that included at least one knee joint movement parameter (moments, kinematics, or muscle activity) and activity limitation or pain outcomes at two time points. Studies were only included if they provided data with sufficient detail, as outlined in section 3.3.5.

We used the MeSH definition of exercise therapy: "A regimen or plan of physical activities designed and prescribed for specific therapeutic goals. Its purpose is to restore normal musculoskeletal function or to reduce pain caused by diseases or injuries" (National Centre for Biotechnology Information, 2021). Studies that assessed the effects of orthotics, shoes, braces, weight loss, or a single session of exercise were excluded. Due to a lack of translation resources, only studies written in English were included.

Abstracts, then full texts (as appropriate), were independently screened for eligibility by two reviewers (J.T. and E.T.). All disagreements were resolved by discussion between reviewers until consensus.

### 3.3.3 Risk of Bias Within Studies

We aimed to investigate the relationship between changes in two outcome measures. As most studies did not investigate that relationship directly, we instead assessed for measuring change in outcome.

We assessed the thoroughness of description of the population, the reliability and validity of outcome measures, data missingness, appropriateness of data analysis (correlation/regression or within-group mean change), and whether the assessor was blinded to the results of both the movement and clinical outcome measures. We adapted the Joanna Briggs Institute critical appraisal checklist for cohort studies (**Appendix 3-2**) (Moola et al., 2017). The original version of this checklist has been peer reviewed, and an adapted version of this tool has been previously used for a similar purpose (Wernli, Tan, et al., 2020). Two reviewers independently assessed the quality of each article included in this review, and all conflicts were resolved by discussion. The percentage agreement between the two reviewers was 91% to 100% (median, 100%; interquartile range, 95% to 100%).

Studies were at high risk of bias when at least one of the most critical sources of bias was graded negatively. Studies were at overall low risk of bias only when all items were judged positively. Critical sources of bias related to reliability and validity in measuring outcomes, statistical analysis, and whether the assessor was blinded to both outcomes.

### 3.3.4 Certainty of Evidence Across Studies

To rate the certainty of evidence across studies, we used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool. All studies were analysed as cohort studies, because treatment allocation was irrelevant to our research question. We assessed the GRADE domains of risk of bias, imprecision, inconsistency, indirectness, and publication bias in relation to our research question about whether change occurred in movement parameters, activity limitations, and pain. For details about GRADE ratings and how the tool was applied, see **Appendix 3-3**.

### 3.3.5 Data Analysis

#### Data Extraction

A data-extraction table was adapted from a previous systematic review (Wernli, Tan, et al., 2020). The reviewers (J.T. and E.T.) independently extracted study design, sample size, participant characteristics, intervention (type, dose, duration), method of assessment, and prescore, postscore, and change score of movement parameter, activity limitation, and pain outcomes.

#### Outcomes of Interest

Studies were included if they investigated both a knee-specific movement parameter during a functional activity and a clinical outcome measured before and after the intervention. Movement parameters were required to be objective, instrument-assisted measures of knee joint movement during a functional activity (e.g. gait, negotiating stairs, or sit-to-stand). The movement parameters of interest were independent measures of knee joint (a) moments (e.g. knee adduction moment), (b) kinematics (e.g. knee flexion range of movement), and (c) muscle activity (e.g. muscle co-contraction ratio). Movement parameters not measured during a functional activity (e.g. supine knee flexion/extension or dynamometry) or combined measures of movement (e.g. via factor analysis) were excluded.

Clinical outcomes included activity limitation or pain rated with self-reported questionnaires (e.g. the Western Ontario and McMaster Universities Osteoarthritis

Index (WOMAC) physical function and/or pain subscales or the visual analogue scale for pain).

**Data Synthesis**

With data extraction from randomised controlled trials, we considered only change within the exercise intervention arm. If a study included two exercise intervention arms, both were included as separate cohorts.

To explore the relationship between changes in movement parameter and changes in activity limitation or pain, we provide a descriptive synthesis of studies reporting tests of association between the changes in movement and changes in activity limitation or pain. Because only two studies calculated the association directly by correlation analysis (DeVita et al., 2018; Preece et al., 2016) we also report the frequency and direction of change for all outcomes for each co-occurrence that could be tested.

If mean change was not reported, change was calculated by subtracting the postscore from the prescore. When the 95% confidence interval (CI) of the change was not reported, it was calculated from the change score variance. Missing SDs of change scores (SDchange) were estimated from pre and post SDs, using the recommended method for "imputing a change-from-baseline standard deviation using a correlation coefficient" in chapter 6.5.2.8 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins et al., 2020). The change SD was not reported in eight studies (DeVita et al., 2018; Fisher et al., 1997; Foroughi et al., 2011a; Foroughi et al., 2011; Gaudreault et al., 2011; Shen et al., 2008; Shull, Silder, et al., 2013; Wang et al., 2016) and thus was estimated from prescore and postscore values. The standardised response mean (SRM) was calculated (change mean/SDchange), which is the within-group effect size used for repeated measures (Middel & van Sonderen, 2002). The SRM was standardised for dependent samples, allowing for use of Cohen's threshold for effect sizes within a single group (Middel & van Sonderen, 2002). We defined the presence of change as a prescore-minus-postscore change score with a 95% CI that did not cross zero and an SRM greater than 0.2 (at least a small effect size) (Middel & van Sonderen, 2002). Therefore, studies that reported medians were omitted.

### 3.3.6 Deviations from the PROSPERO Registration

This review differs from the PROSPERO registration in that we omitted the question on the magnitude of change when a relationship was present, due to a lack of studies that directly investigated the change relationship at an individual person-level (such as using correlation analyses). We include data for the reader about changes in outcome for all outcome measures in the online appendices.

## 3.4 Results

The search strategy identified 3182 potentially suitable articles, and 22 met the inclusion criteria (n = 936 participants) (Bennell et al., 2010; Bennell et al., 2014; Brenneman et al., 2015; Cheung et al., 2018; DeVita et al., 2018; Fisher et al., 1997; Foroughi et al., 2011a; Foroughi et al., 2011; Gaudreault et al., 2011; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Hunt & Takacs, 2014; King et al., 2008; Lim et al., 2008; Preece et al., 2016; Roper et al., 2013; Shen et al., 2008; Shull, Silder, et al., 2013; Sled et al., 2010; Turcot et al., 2009; Wang et al., 2016; Zhu et al., 2016). The PRISMA flow diagram is presented in Error! Reference source not found. and the characteristics of the 22 articles (which included a total of 26 intervention arms) in **Appendix 3-4**. Two studies (Foroughi et al., 2011a; Foroughi et al., 2011) reported data from the same cohort; participants were counted once (Foroughi et al., 2011a), and duplicate data for any movement parameter were included once (Foroughi et al., 2011a). Heterogeneity of movement parameters and interventions precluded meta-analysis.

### 3.4.1 Risk of Bias Within Studies

One study was at low risk of bias, one was at medium risk, and 20 were at high risk (**Table 3-1**). The most common reasons for a lower rating were uncertainty about whether the assessor was blinded to both movement parameters and clinical outcomes at both time points (unclear in 18/22 studies) and missing data (not adequate or unclear in 6/22 studies).

**Figure 3-1. PRISMA flow chart.**

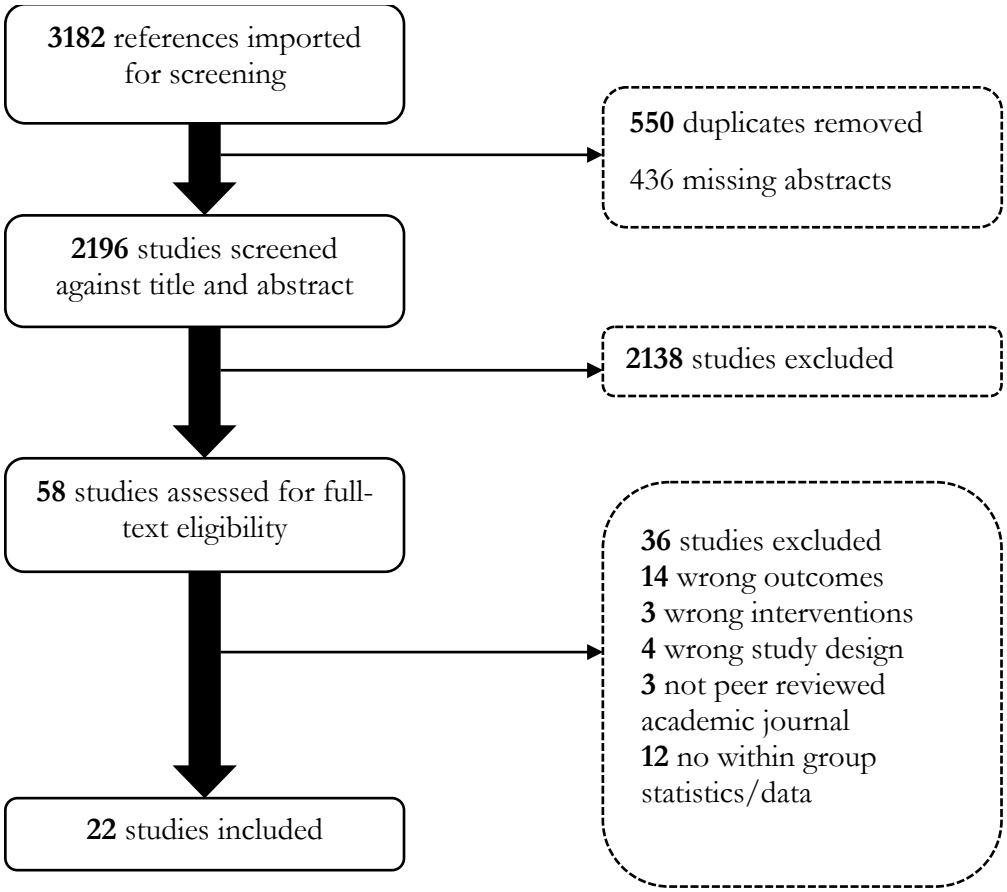


**Table 3-1. Risk-of-bias assessment**

| Author, year | 1. Demographics | 2. Selection criteria | 3. Missingness | 4. Movement standard/valid* | 5. Movement reliable* | 6. Pain/Act Lim standard/valid* | 7. Pain/Activity Limitation reliable* | 8. Appropriate analysis* | 9. Assessor blinding* | Overall Risk of Bias (low, medium, high) |
|---|---|---|---|---|---|---|---|---|---|---|
| (Bennell et al., 2010) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Bennell et al., 2014) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Brenneman et al., 2015) | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | High |
| (Cheung et al., 2018) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (DeVita et al., 2018) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Fisher et al., 1997) | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Foroughi et al., 2011a) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | Low |
| (Foroughi et al., 2011) | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | Medium |
| (Gaudreault et al., 2011) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Holsgaard-Larsen et al., 2017) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Hunt et al., 2018) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Hunt & Takacs, 2014) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (King et al., 2008) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Lim et al., 2008) | 1 | 1 | U | 1 | 1 | 1 | 1 | 1 | U | High |
| (Preece et al., 2016) | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Roper et al., 2013) | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Shen et al., 2008) | 1 | 1 | U | 1 | 1 | 1 | 1 | 1 | U | High |
| (Shull, Silder, et al., 2013) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Sled et al., 2010) | 1 | 1 | U | 1 | 1 | 1 | 1 | 1 | U | High |
| (Turcot et al., 2009) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | High |
| (Wang et al., 2016) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| (Zhu et al., 2016) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | U | High |
| Domain level prevalence | 91% | 100% | 73% | 100% | 100% | 100% | 100% | 100% | 9% | |

Abbreviations: *= critical risk of bias, 1 = Adequate, 0 = Not adequate, U = Unclear.
High risk = any of the critical risks scored negatively or were unclear; Medium risk = all of the critical risks scored positively but some non-critical risks scored negatively or were unclear; Low risk = scored positively on all items.

### 3.4.2 Certainty of Evidence Across Studies

There was an overall low certainty of evidence for changes in outcomes and movement parameters (Error! Reference source not found.). We downgraded the starting position of "medium quality" to "low quality" and have limited confidence in the effect estimate due to the reasons detailed in **Appendix 3-2**.

**Table 3-2. GRADE summary of findings for each outcome**

| Change outcome | Studies (n) | Risk of Bias | Imprecision | Inconsistency | Indirectness | Publication Bias | GRADE Rating |
|---|---|---|---|---|---|---|---|
| Movement parameters | 22 | Very serious | Not serious | Very serious | Very serious | Likely | Very low |
| WOMAC PF subscale | 15 | Very serious | Not serious | Not serious | Not serious | Not likely | Low |
| KOOS ADL subscale | 3 | Very serious | Not serious | Not serious | Not serious | Not likely | Low |
| WOMAC pain subscale | 16 | Very serious | Not serious | Not serious | Not serious | Not likely | Low |
| KOOS pain subscale | 3 | Very serious | Not serious | Not serious | Not serious | Not likely | Low |
| Pain VAS/NRS | 8 | Very serious | Not serious | Not serious | Not serious | Not likely | Low |
| FSI difficulty subscale | 1 | Very serious | NA | NA | NA | NA | NA |
| FSI pain subscale | 1 | Very serious | NA | NA | NA | NA | NA |

Abbreviations: ADL, activities of daily living; FSI, Functional Status Index; GRADE, Grading of Recommendations Assessment, Development and Evaluation; KOOS, Knee injury and Osteoarthritis Outcome Score; NA, not applicable; NRS, numeric rating scale; PF, physical function; VAS, visual analogue scale; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

### 3.4.3 Synthesis of Results

Two studies directly investigated the change relationship via correlation analysis. Gait was the only functional activity studied (22 studies, 26 intervention arms): there were eight knee joint moment, 15 kinematic, and four muscle activity parameters investigated. The within-group mean prescore, postscore, and change score (with 95% CI) and the SRM for movement parameters for each study are detailed in **Appendix 3-5**, and for clinical outcomes in **Appendix 3-6**. The number of occasions a change in movement parameter was investigated and the frequency (%) at which it occurred are reported in **Table 3-3.**

A within-group change (an SRM greater than 0.2, with the 95% CI of the mean change not crossing zero) in any gait- related movement parameter occurred 26% of the time (22/84), a within-group change in gait-related moments occurred 27% of the time (15/56) (Bennell et al., 2010; Cheung et al., 2018; DeVita et al., 2018; Foroughi et al., 2011a; Gaudreault et al., 2011; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Hunt & Takacs, 2014; Lim et al., 2008; Shull, Silder, et al., 2013; Wang et al., 2016), and a within-group change in gait-related kinematics occurred 25% of the time (7/28) (Roper et al., 2013; Turcot et al., 2009; Wang et al., 2016; Zhu et al., 2016). There were no within-group change data available for any muscle activity parameter.

**Table 3-3. Changes in knee joint movement parameters during gait**

| Movement parameter | Occasions investigated (n) | Change occurred n(%) |
|---|---|---|
| Moments | | |
| 1st Peak knee adduction moment* | 20 | 7 (35%) |
| Knee flexion moment | 14 | 0 |
| Peak knee adduction moment impulse | 8 | 3 (37%) |
| Knee extension moment | 5 | 2 (40%) |
| 2nd Peak knee adduction moment | 5 | 2 (40%) |
| Peak knee internal rotation moment | 1 | 0 |
| Peak knee abduction moment | 2 | 1 (50%) |
| Peak knee external rotation moment | 1 | 0 |
| | 56 | 15 (27%) |
| Kinematics | | |
| Knee flexion range | 2 | 0 |
| Peak knee flexion angle early stance | 4 | 1 (25%) |
| Peak knee adduction angle | 5 | 0 |
| Peak knee abduction angle | 3 | 2 (67%) |
| Max sagittal angular velocity stance | 2 | 1 (50%) |
| Min transverse angular velocity swing | 2 | 1 (50%) |
| Max sagittal angular velocity swing | 2 | 0 |
| Peak knee external rotation angle | 1 | 0 |
| Peak knee flexion angle swing | 1 | 0 |
| Peak knee internal rotation angle | 1 | 0 |
| Acceleration anterior-posterior | 1 | 1 (100%) |
| Knee angle initial contact | 1 | 1 (100%) |
| Acceleration medial-lateral | 1 | 0 |
| Adduction-abduction range | 1 | 0 |
| Internal-external rotation range | 1 | 0 |
| | 28 | 7 (25%) |

Change was defined as occurring when the 95% confidence interval did not cross zero and when the standardised response mean was greater than 0.2.
*The study by (Foroughi et al., 2011a) was excluded from the count because data were reported in (Foroughi et al., 2011)

The most frequently studied movement parameter was first peak knee adduction moment, which was also the one that changed most frequently (7/20, 35%). An improvement in activity limitation or pain occurred 90% of the time (139/155) a cooccurrence of change was tested against gait parameters. No study reported worsening of activity limitation or pain after intervention.

### 3.4.4 Studies Testing a Correlation Between Gait Parameters and Clinical Outcomes

One study (DeVita et al., 2018) directly investigated the association between change in two movement parameters (peak flexion angle at early stance and peak knee

extension moment) and change in activity limitation. There was no correlation between change in any movement parameter and change in activity limitation. Two studies directly investigated the association between change in six movement parameters (precontact and early stance medial and lateral hamstrings-quadriceps co-contraction ratios (Preece et al., 2016), peak flexion angle at early stance and peak knee extension moment (DeVita et al., 2018)) and change in pain. Change in one muscle activity parameter (precontact medial hamstrings-quadriceps co-contraction ratio) had a statistically significant correlation with a change in WOMAC pain subscale score ($r = 0.45$, $p < 0.05$) (Preece et al., 2016).

### 3.4.5 Co-occurrence of Within-Group Mean Change Between Gait Parameters and Clinical Outcomes

A co-occurrence of change between gait-related movement parameters and activity limitation or pain was uncommon. A change (SRM greater than 0.2, with the 95% CI of the mean change not crossing zero) in both a movement parameter and a clinical outcome occurred 24.5% of the times (38/155) a comparison could be made (Bennell et al., 2010; Cheung et al., 2018; DeVita et al., 2018; Foroughi et al., 2011a; Gaudreault et al., 2011; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Hunt & Takacs, 2014; Lim et al., 2008; Roper et al., 2013; Shull, Silder, et al., 2013; Turcot et al., 2009; Wang et al., 2016; Zhu et al., 2016). Those instances where there was within-group change in movement parameters are reported in **Table 3-4,** cross-tabulated with change in activity limitation or pain. **Appendix 3-7** expands on **Table 3-4**, presenting results for those instances where movement parameters did not demonstrate change. Where co-occurrence of change was observed between first peak knee adduction moment (Bennell et al., 2010; Cheung et al., 2018; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Lim et al., 2008; Shull, Silder, et al., 2013) or peak knee adduction moment impulse (Gaudreault et al., 2011; Holsgaard-Larsen et al., 2017; Hunt et al., 2018) and a clinical outcome, the direction of change for the movement parameter was variable (**Table 3-4**).

**Table 3-4. Direction of change in gait-related movement parameters, activity limitation and pain for movement parameters that demonstrated a change**

| Movement Parameter | | Activity Limitation | | | Pain | | |
|---|---|---|---|---|---|---|---|
| | | Improved | No Change | Worse | Improved | No Change | Worse |
| **Moments** | | | | | | | |
| 1st Peak knee adduction moment | Increased | 4 (21%) | 1 (5%) | 0 | 4 (20%) | 1 (5%) | 0 |
| | No Change | 9 (47%) | 3 (16%) | 0 | 10 (50%) | 3 (15%) | 0 |
| | Decreased | 2 (11%) | 0 | 0 | 2 (10%) | 0 | 0 |
| Peak knee adduction moment impulse | Increased | 1 (14%) | 0 | 0 | 1 (12%) | 0 | 0 |
| | No Change | 4 (57%) | 0 | 0 | 5 (63%) | 0 | 0 |
| | Decreased | 2 (29%) | 0 | 0 | 2 (25%) | 0 | 0 |
| Knee extension moment | Increased | 1 (20%) | 0 | 0 | 1 (20%) | 0 | 0 |
| | No Change | 2 (40%) | 1 (20%) | 0 | 3 (60%) | 0 | 0 |
| | Decreased | 1 (20%) | 0 | 0 | 1 (20%) | 0 | 0 |
| 2nd Peak knee adduction moment | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 3 (75%) | 0 | 0 | 3 (60%) | 0 | 0 |
| | Decreased | 1 (25%) | 0 | 0 | 2 (40%) | 0 | 0 |
| Peak knee abduction moment | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 (50%) | 0 | 0 | 1 (50%) | 0 | 0 |
| | Decreased | 1 (50%) | 0 | 0 | 1 (50%) | 0 | 0 |
| **Kinematics** | | | | | | | |
| Peak knee flexion angle early stance | Increased | 1 (25%) | 0 | 0 | 1 (25%) | 0 | 0 |
| | No Change | 2 (50%) | 1 (25%) | 0 | 3 (75%) | **0** | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee abduction angle | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 (33%) | 0 | 0 | 1 (33%) | 0 | 0 |
| | Decreased | 2 (67%) | 0 | 0 | 2 (67%) | 0 | 0 |
| Max sagittal angular velocity stance | Increased | 0 | 0 | 0 | 1 (50%) | 0 | 0 |
| | No Change | 0 | 0 | 0 | 0 | 1 (50%) | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Min transverse angular velocity swing | Increased | 0 | 0 | 0 | 1 (50%) | 0 | 0 |
| | No Change | 0 | 0 | 0 | 0 | 1 (50%) | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Acceleration anterior-posterior | Increased | 0 | 0 | 0 | 1 (100%) | 0 | 0 |
| | No Change | 0 | 0 | 0 | 0 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Knee angle at initial contact | Increased | 1 (100%) | 0 | 0 | 1 (100%) | 0 | 0 |
| | No Change | 0 | 0 | 0 | 0 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |

Change was defined as 95% confidence interval does not cross zero and standardised response mean >0.2. * Vote counting system – 1.0 vote per parameter. Where multiple measures were used for an outcome, the 1 vote was divided by the total number times used. For example, if pain was reported using WOMAC pain and NRS pain, then the vote was divided equally, therefore if WOMAC pain changed and NRS pain did not change the score was 0.5. See

**Appendix** 3-7. Direction of a change in movement parameter, activity limitation and pain (count (%)) for movement parameters that did not demonstrate a change.

for movement parameters that did not demonstrate a change.

For gait, the co-occurrence of change was observed between moment parameters and activity limitation 26% of the time (13/50) (Bennell et al., 2010; Cheung et al., 2018; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Lim et al., 2008; Shull, Silder, et al., 2013) and between kinematic parameters and activity limitation 18% of

the time (4/22) (Wang et al., 2016; Zhu et al., 2016). There were no studies that provided sufficient data to establish a co-occurrence of change between muscle activity and activity limitation.

Co-occurrence of change was observed between moment parameters and pain 26% of the time (14/54) (Bennell et al., 2010; Cheung et al., 2018; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Lim et al., 2008; Shull, Silder, et al., 2013) and between kinematic parameters and pain 24% of the time (7/29) (Roper et al., 2013; Turcot et al., 2009; Wang et al., 2016; Zhu et al., 2016). There were no studies that provided sufficient data to establish a co-occurrence of change between muscle activity and pain.

## 3.5 Discussion

We aimed to quantify the relationship between change in knee joint movement parameters and change in activity limitation or pain after an exercise intervention for people with knee osteoarthritis. The included studies were mostly of low quality and exclusively focused on gait, overlooking other important functional tasks relevant to people with knee osteoarthritis (e.g. sit-to-stand or ascending stairs). Gait parameters were predominantly focused on knee adduction moment, with limited investigation of kinematic and muscle activity parameters.

Only two studies directly assessed a relationship via correlation analysis (DeVita et al., 2018; Preece et al., 2016). Change in only one movement parameter of the six investigated had a significant (moderate) correlation with improvement in pain (Preece et al., 2016). We found a within-group co-occurrence of change between gait-related movement parameters and activity limitation or pain to be infrequent (24.5% of the time), even though clinical outcomes improved frequently (90% of the time).

### 3.5.1 Interpretation of Findings

Our results align with previous reviews across multiple body regions that report infrequent (Laird et al., 2012; Wernli, Tan, et al., 2020) or absent (Ferreira et al., 2015; Nodehi Moghadam et al., 2020; Richards et al., 2017) relationships between changes in movement patterns and clinical outcomes. We expanded on previous reviews in people with knee osteoarthritis (Ferreira et al., 2015; Richards et al., 2017) by

including a broader range of movement parameters and interventions (including active intervention control groups) and by reporting within-group change and additional results from cohort studies.

There were only two studies (DeVita et al., 2018; Preece et al., 2016) that directly investigated the strength of the relationship between changes in movement parameters (peak flexion angle at early stance, peak knee extension moment, and medial and lateral hamstrings-quadriceps co-contraction ratio at precontact and early stance) and changes in activity limitation or pain, using the most precise statistical approaches (correlation or regression). Only one relationship was significant: a (moderate) correlation between a reduction in medial hamstrings-quadriceps co-contraction ratio at precontact and improvement in WOMAC pain subscale score (Preece et al., 2016). In a recent prospective study over one year, higher medial hamstrings-quadriceps co-contraction was associated with faster progression of medial knee osteoarthritis resulting from increased medial loading (Hodges et al., 2016). Considering that knee joint effusion is an independent predictor of weight-bearing pain (Lo et al., 2009), and that people who have knee osteoarthritis and joint effusion have higher co-contraction than those with knee osteoarthritis but without effusion, the preliminary and promising results warrant further investigation (Preece et al., 2016).

Fifteen studies in our review aimed to reduce medial joint loading through a reduction in knee adduction moment, which in turn may help prevent structural progression of the condition. On the other hand, physical therapists commonly prescribe load-modifying interventions for people with knee osteoarthritis, aiming to improve function (Teo et al., 2020). There are a variety of exercise modes (e.g. gait retraining, neuromuscular exercise) commonly prescribed by physical therapists that may modify or normalise medial joint loading. Beyond preventing structural progression, physical therapists have traditionally treated those with specific pathology as a homogeneous group and viewed modifying movement patterns through a kinesiopathological model of care – assuming that modifying a movement pattern to a specific, predetermined "normal" is necessary to reduce pain and improve function (Harris-Hayes et al., 2010; Lehman, 2018). Given that activity limitation or

pain almost always improved, despite a co-occurrence with change in movement parameters only 25% of the time, we suggest that changes in other factors may relate to the clinical benefits observed during exercise interventions.

We found that improvements in clinical outcomes can occur together with a change in movement in a direction that is theoretically "detrimental." Despite elevated first peak knee adduction moment being associated with structural progression of knee osteoarthritis (Bennell et al., 2011; Chehab et al., 2014; Thorp et al., 2006), we found more studies reporting concurrent improvements in clinical outcomes with increased (Bennell et al., 2010; Holsgaard-Larsen et al., 2017; Hunt et al., 2018; Lim et al., 2008), rather than reduced (Cheung et al., 2018; Shull, Silder, et al., 2013), first peak knee adduction moment after exercise (**Table 3-4**). There are theories about why both directions of association may be plausible that may be separate from theories about why exercise is helpful in the absence of a change in movement parameters (Runhaar et al., 2015). Increased knee adduction moment is a risk factor for structural progression of medial knee osteoarthritis (Chehab et al., 2014), but the association between knee adduction moment and pain differs according to underlying radiological severity (Hall et al., 2017). There may be a balancing act for clinicians: helping to manage pain or improve function on one hand and preventing structural progression on the other.

Improvements in clinical outcomes usually occurred with no change in a movement parameter. Therefore, it is unlikely that normalising movement patterns is a prerequisite for a change in clinical outcomes, and different exercise interventions may act through different mechanisms (Runhaar et al., 2015). In people with knee osteoarthritis, exercise has the potential to impact risk factors for poor outcome, such as obesity (Hall et al., 2019), muscle strength (Dekker et al., 2009; Runhaar et al., 2015), sedentary lifestyle (Alentorn-Geli et al., 2017), and psychological factors such as depression and low self-efficacy (Briani et al., 2018; van Dijk et al., 2006). Change in those risk factors may have a stronger relationship with clinical outcomes than normalising movement patterns.

Group-level change in gait-related movement parameters occurred infrequently and bidirectionally in the limited number of studies where change was observed.

Therefore, exercise may not predictably or consistently change movement parameters in a specific direction across participants. One reason is that the populations studied may not have had the potential to change. For example, participants may not have had activity limitation or pain related to gait or impairments of the investigated movement parameter at baseline. Also, some participants may have had fixed deformity, in which exercise was unable to change the way they move during gait. It is also possible that the intervention did not target the impairment by using an exercise that reflects the specific functional activity (e.g. gait retraining versus neuromuscular exercise).

### 3.5.2 Strengths

Our review describes all movement parameters that have been investigated for change alongside change in activity limitation or pain. We provide the first estimation of the co-occurrence of changes in diverse movement parameters during functional activities and changes in activity limitation or pain in people with knee osteoarthritis.

### 3.5.3 Limitations

There is a risk of selective reporting bias (most studies were not prospectively registered), language bias (only English-language publications were included), and publication bias (only peer-reviewed studies were included). As studies did not always include mean ± SD change, we estimated those values (using a best-practice approach as described by the imprecision. Some other kinetic parameters not considered in our review (e.g. force) may have stronger relationships (DeVita et al., 2018). Because we were interested only in a relationship between change in outcomes at the within-group level, it is not known whether the observed change was the result of the exercise intervention, which would need to be investigated in randomised controlled trial designs. Finally, the estimates of a change relationship were mostly described in the included studies at the group level, which provides limited ability to make inferences about relationships at the individual-level.

### 3.5.4 Clinical Implications

The results of this systematic review do not appear to support the targeting of a change in movement patterns during a functional activity to improve activity limitation or pain in people with knee osteoarthritis. In the evidence that was

available, gait-related knee joint movement patterns changed infrequently across studies, and these changes in movement were infrequently related to improvement in the clinical outcomes of activity limitation and pain at a group level. However, our confidence in this conclusion is limited due to the low-quality evidence.

The pain experience of a person with knee osteoarthritis is multifactorial (Georgiev & Angelov, 2019). The current evidence supports education, weight loss, and exercise therapy for improving clinical outcomes (Hall et al., 2018; Mihalko et al., 2019). Exercise is strongly advocated and well supported by evidence (Bannuru et al., 2019; Fransen et al., 2015) but it remains unclear what role changing movement may have in directly improving activity limitation and pain in a person with knee osteoarthritis. In people with low back pain, there were moderate to large correlations between changes in movement parameters and activity limitations after an intervention targeting individually relevant movements (Wernli, O'Sullivan, et al., 2020). However, individually relevant movement parameters and activity limitations were diverse and specific to the individual (Wernli, O'Sullivan, et al., 2020). These findings in people with low back pain may suggest similar diversity and the need to provide individualised care for people who have knee osteoarthritis or other conditions, although that has yet to be tested.

### 3.5.5 Unanswered Questions and Future Research

There is an absence of high-quality research, with appropriate assessor blinding, investigating the relationship between changes in how a person with knee osteoarthritis moves during functional activities and improvement in clinical outcomes. Researchers might consider two methodological approaches to quantify a relationship between changes in outcomes: (a) apply correlation or regression analyses of change scores to investigate person-level change relationships, or (b) consider meta-regression of randomised controlled trial data (comparative treatment effects) to determine whether changes are the result of the intervention and to quantify the association between those changes, ideally using participant-level data.

To date, research has focused only on gait. It is unclear whether movement parameters during other functional activities (e.g. ascending stairs or standing from a chair) change after exercise intervention, and whether there is a relationship with

change in clinical outcomes. Further research may reveal stronger change relationships between individually targeted activities and movement parameters and clinical outcomes (Wernli, O'Sullivan, et al., 2020).

Other movement parameters, such as the hamstrings-quadriceps co-contraction ratio, warrant further investigation due to the demonstrated link with pain intensity (Preece et al., 2016). Future studies could also investigate people with knee osteoarthritis who demonstrate symptom relief following intervention, to determine whether change in knee adduction moment is protective or deleterious for long-term structural severity, activity limitation, or pain.

## 3.6 Conclusion

Targeting a change in movement patterns is unlikely to improve activity limitation or pain in people with knee osteoarthritis. Gait movement parameters have been most frequently explored and change infrequently after an exercise intervention, despite most studies reporting improvements in activity limitation and pain.

## 3.7 Key Points

Findings: A relationship between group-level change in gait-related movement parameters and change in activity limitation or pain occurred only 24.5% of the time comparisons could be made, despite improvements in clinical outcomes occurring 90% of the time, across low-quality studies.

Implications: It is unlikely that improvements in clinical outcomes after exercise are related to changes in gait-related knee joint moments, kinematics, and muscle activity at the group level. Clinicians should prescribe exercise interventions for people with knee osteoarthritis, while being mindful that these findings suggest that changing movement patterns may be unrelated to clinical outcomes.

Caution: The majority of studies were of low quality and did not directly investigate an association between changes in movement parameters and clinical outcomes at an individual level. Other activities besides gait have not yet been investigated.

## 3.8 Study Details

Author Contributions: All authors contributed to the conception, design, analysis, and interpretation of this study. All authors contributed to drafting and critically revising the manuscript and assume responsibility for the work.

Data Sharing: All data relevant to the manuscript are included in the article or within the online appendices.

Patient and Public Involvement: There was no patient or public involvement in this research.

Acknowledgements: The authors would like to acknowledge Senior Faculty Librarian Diana Blackwood for her contribution to the electronic search strategy for this manuscript.

## 3.9 Appendices

## Appendix 3-1. Search strategy (Medline/EMBASE)

| | | |
|---|---|---|
| 1 | Diagnosis | osteoarthritis, knee.sh. OR (knee and osteoarthritis).ti,ab. OR (knee and arthritis).ti,ab. OR (knee and degenerative and joint).ti,ab. OR knee osteoarthritides.ti,ab. OR knee OA.ti,ab. OR osteoarthritides, knee.ti,ab. OR knee, OA of.ti,ab. OR knees, osteoarthritis of.ti,ab. OR osteoarthritis of knee*.ti,ab. OR OA of knee*.ti,ab. |
| 2 | Intervention | Exp rehabilitation/ OR physical therapy modalities/ OR physiotherap*.af. OR physical therap*.af. OR<br><br>Exp Exercise therapy/ OR exercise therapy.af. OR Exercise movement techniques/ OR exercise*.af. OR therapeutic exercise.af. OR functional exercise.af. OR resistance training/ OR strength*.af. OR plyometric exercise/ OR closed kinetic chain exercise*.af. OR open kinetic chain exercise*.af. OR group exercise.af. OR<br><br>Gait $training.af. OR<br><br>biofeedback.af. OR bio-feedback.af. OR (neuromuscular rehabilitation).af. OR (EMG biofeedback).af. OR retraining.af. OR neuromuscular facilitation.af. OR<br><br>hydrotherapy.af. OR yoga/ OR yoga.af. OR tai ji/ OR tai chi.af. OR dance therapy/ OR aerobic exercise.af. |
| 3 | Pain | Pain measurement/ OR pain.ti,ab. |
| 4 | | (VAS OR visual analogue scale OR NRS OR Numeric* pain rating scale OR ICOAP OR (Intermittent and Constant Osteoarthritis Pain) OR MPQ OR McGill Pain Questionnaire).af. |
| 5 | | 3 OR 4 |
| 6 | Activity limitation | Disability evaluation/ OR Activities of daily living/ OR (function* OR functional outcome OR disability OR activity limitation).af. |
| 7 | | (WOMAC OR (Western Ontario and McMaster Universities Osteoarthritis Index) OR KOOS OR (Knee Injury and Osteoarthritis Outcome Score) OR HAQ OR (Health Assessment Questionnaire) OR PROMIS OR (Patient Reported Outcomes Measurement Information System) OR PSFS OR (Patient Specific Functional Scale)).af. |
| 8 | | 6 OR 7 |
| 9 | | patient reported outcome measures/ OR Self report/ OR treatment outcome/ OR patient outcome assessment/ |
| 10 | Primary outcome | 5 OR 8 OR 9 |
| 11 | Movement outcome | exp biomechanical phenomena/ OR biomechanic*.af. OR movement pattern*.af. OR exp biomechanical/ OR Gait analysis/ |
| 12 | | kinematic*.af. OR angle.af. OR angular velocity.af. |
| 13 | | Exp Kinetics/ OR kinetic*.af. OR moment.af. OR torque/ or impulse.af. |
| 14 | | exp electromyography/ OR EMG.af. OR sEMG.af. OR motor 97ctive*.af. OR 97ctive* pattern*.af. OR muscle 97ctive*.af. OR co-contraction.af. OR co-contraction index.af. |
| 15 | | 11 OR 12 OR 13 OR 14 |
| 16 | Functional activity | human activities/ OR musculoskeletal physiological phenomena/ |
| 17 | | Gait/ OR locomotion/ OR walking/ OR gait.af. OR walk*.af. |
| 18 | | Sit* to stan*.af. OR stan* to sit*.af. OR rising.af. |
| 19 | | stair climbing/ OR stair*.af. OR step*.af. OR stair climbing.af. |
| 20 | | 16 OR 17 OR 18 OR 19 |
| 21 | Movement during functional activity | 15 and 20 |
| 22 | Trials | randomized controlled trial.pt. OR controlled clinical trial.pt. OR clinical trial.pt. OR random*.ti,ab. OR ((doubl* OR singl*) and blind*). Ti,ab. OR exp clinical trial/ OR crossover. Ti,ab. OR clin* trial.ti,ab. OR (control* and (trial* OR stud*)).ti,ab. OR ((singl* OR doubl* OR tripl* OR trebl*) and (blind* OR mask*)).ti,ab. OR placebo.ti,ab. OR research design/ OR comparative study/ |
| 23 | | meta-analysis.pt. OR (meta-anal* OR metaanaly* OR meta analy*).ti,ab. |
| 24 | Cohorts | Case-control studies/ OR cohort studies/ OR comparative study/ OR evaluation studies/ OR follow-up studies/ OR longitudinal studies/ OR prospective studies/ OR treatment outcome/ OR exp Clinical Study/ OR observational study/ |
| 25 | Trial type | 22 OR 23 OR 24 |
| 26 | | 1 and 2 and 10 and 21 and 25 |
| 27 | | (neoplasm* OR cancer*).ti. |
| 28 | Not | Cross-Sectional Studies/ OR cross-sectional.pt. OR cross sectional.pt. OR reliability.ti. |
| 29 | | 28 OR 29 |
| 30 | | 26 not 29 |
| 31 | | Limit 31 to English language |

## Appendix 3-2. Adapted Joanna Briggs Institute Critical Appraisal Tool for Cohort Studies.

| Risk of Bias Item | Description |
|---|---|
| 1. Were the **demographic** and **clinical characteristics** of the participants adequately described? | Description of the participants adequate enough to understand how generalisable (externally valid) the findings might be to other settings. |
| 2. Were the **selection criteria** for participants adequately described? | Knowledge of how the participants were selected is important to also understand how generalisable the findings are. For example, if the sample only included participants with certain traits (such as only including people with generalised hypermobility). |
| 3. Was there an acceptable **missingness of data**? | Data missingness is not in excess of 20% (>80% of the data analysed). If applicable, statistical analysis that handles missing data well is used (for example: multiple imputation, Generalised Estimating Equations, Mixed Effect models, Latent Growth models). Non-random missingness from people in a cohort can reduce external validity, while missingness of data timepoints in longitudinal studies can reduce internal validity. |
| 4. Were the **movements** assessed in a **standardised** and **valid** way? | The study clearly describes the method of measurement of movements. Assessing validity requires that a 'pseudo-gold standard' is available with which the measure has been compared or that construct, and concurrent validity have been quantified in some other way. Measure needs to be well described and referenced unless it is well known. |
| 5. Were **movements assessed** in a **reliable** way? | In this context, reliability (reproducibility) refers to the quantification of test-retest variability. The measure needs to be well known, or if not well known, it needs to be referenced. |
| 6. Were **pain** and/or **activity limitation** assessed in a **standardised** and **valid** way? | The study should clearly describe the method of measurement of pain and/or activity limitation. Validity requires some demonstration (for example a reference) of construct and concurrent validity. |
| 7. Were **pain** and/or **activity limitation** assessed in a **reliable** way? | Reliability refers here to the quantification of test-retest variability. The method of measurement needs to be well known, or if not well known it needs to be referenced. |
| 8. Was appropriate **statistical analysis** used? | Transparency and appropriate selection of the analytical strategy used. In this case, appropriate statistical approaches would include test to determine whether change over time occurred (e.g. T-tests, Mann-Whitney-U / Wilcoxon rank-sum tests, or similar), or they provide sufficient data that such a test can be performed post-hoc. |
| 9. Was the assessment of movement (and its change) **blind** to the **assessment** of pain or activity limitation (and its change) or vice-versa? | Assessor knowledge of participant self-report of pain or activity limitation (or its change) may have biased the assessor's subsequent assessment of movement (or vice-versa). If assessor blinding is reported, a '1 - yes' is scored; if the non-blinding of assessors is reported, a '0 – no' is scored; and if it is not clear if blinding of assessors took place, a 'U – Unclear' is scored. |

Each item scored: 'Yes, No, Unclear or Not Applicable'.

## Appendix 3-3. Scores and reasoning for the five GRADE domains assessing quality of evidence across studies for change in outcomes.

The GRADE guidelines suggest starting at 'low-quality' for research using cohort designs. Because cohort studies are not inferior to randomised controlled trials for assessing within group change, we started with a rating of 'medium quality' and we used GRADE criteria to upgrade or downgrade the quality score.

*Domain 1. Risk of bias* - All but two studies were rated as high risk of bias for investigating change in outcomes in this review. We therefore downgraded to rating of very low-quality was due to lack certainty if assessors of the change in movement parameters were blinded the change in the clinical outcome measures (pain and activity limitation).

*Domain 2. Imprecision* - From 22 studies, 936 individuals with knee osteoarthritis were investigated increasing the generalisability of the findings in this review, albeit that there was diversity in the movement parameters assessed and measurement of clinical outcomes. We have increased confidence in the estimates of change for clinical outcomes with consistent point estimates and confidence intervals across studies. Movement parameters were investigated using measurement equipment with known error within and between sessions (McGinley et al., 2009) increasing the confidence of identifying true change. However, a boundary of the confidence intervals for a change in movement was often close to 0, indicating a possibility that true change did not occur.

*Domain 3. Inconsistency* - We have confidence of a consistent and true estimate of clinical benefit (activity limitation and pain) across the majority of studies, with 90% reporting moderate to large positive effect sizes after exercise interventions and appreciable overlap of confidence intervals. This contrasts with movement parameters, where there were inconsistent point estimates of change with a wide range of confidence intervals (even for the same movement parameter), which reduces confidence of these findings.

*Domain 4. Indirectness* - The population sample, clinical outcome measures (activity limitation and pain) and interventions were representative of those in clinical practice, improving the generalisability of these findings. Similarly, the equipment used to investigate change in movement parameters was quite precise, notwithstanding that it is not widely accessible to clinicians in a clinical environment.

*Domain 5. Publication bias* - The majority of included studies reported no group-level change in movement and this finding did not differ between studies with smaller or larger samples, minimising the likelihood of overestimation of the effect. There may be an overestimation of the frequency across studies where a change in movement parameter occurred due to the potential of selective publication of studies that showed positive findings.

Criteria which could influence a decision to upgrade the quality rating:
*1. Large effect* - Not applicable to this systematic review.
*2. Dose Response* - Not applicable to this systematic review.
*3. Accounted for all plausible residual confounding* - Not applicable to this systematic review.

*Summary*
Starting with a GRADE of 'medium-quality' for cohort derived data, after consideration of all GRADE criteria, we downgraded the classification to low-quality evidence: defined by GRADE as "our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect". The main reason was because, for our research question about a relationship of change between clinical attributes, our judgement is that the degree of indirectness in the included studies and the possibility of a lack of movement assessor blinding, resulted in a meaningful threat to the estimate of how frequently that relationship occurs.

*References for quality of evidence assessment:*
Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336*(7650), 924-926. https://doi.org/10.1136/bmj.39489.470347.ad

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., ... & Guyatt, G. H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology, 64*(4), 401-406. https://doi.org/10.1016/j.jclinepi.2010.07.015

Siemieniuk R, Guyatt G. What is GRADE? *Learn Evidence Based Medicine: Toolkit* 2020; https://bestpractice.bmj.com/info/us/toolkit/learn-ebm/what-is-grade/. Accessed 30 September, 2020.

## Appendix 3-4. Characteristics of included studies.

| Author, year | N (recruited),(% Female) | Diagnostic criteria | Intervention | Study population (BMI, Grade (Kellgren-Lawrence), Age) | Follow-up | Movement measurement device | Movement outcome | Patient reported outcome |
|---|---|---|---|---|---|---|---|---|
| (Bennell et al., 2010) | 45, (51 %) | ACR | Neuromuscular exercise | 27.5(4.7),<br>GII 15 GIII 15 GIV 15,<br>64.5(9.1) | 13 weeks | Optical (8 camera Vicon system) with AMTI force plates | Gait<br>1st PKAM<br>PKAM impulse | WOMAC function<br>WOMAC pain |
| (Bennell et al., 2014) | 100, (52 %) | ACR | Neuromuscular exercise<br><br>Quadriceps strengthening | 29.6(3.9)<br>GII 9, GIII 21, GIV 20<br>62.7(7.3),<br><br>29.7(4.3)<br>GII 13, GIII 22, GIV 15,<br>62.2(7.4) | 13 weeks | Optical (12 camera Vicon system) with AMTI force plates | Gait<br>1st PKAM<br>PKAM impulse<br>PKFM | WOMAC function<br>WOMAC pain<br>NRS pain movement<br>NRS pain overall |
| (Brenneman et al., 2015) | 45, (100 %) | ACR | Yoga-inspired strengthening program | Ht 1.63 (0.06), Wt 78.1(14.8),<br>NR,<br>60.3(6.5) | 12 weeks | Optical (9 camera Optotrak system) with AMTI force plates | Gait<br>1st PKAM | KOOS ADL<br>KOOS Pain |
| (Cheung et al., 2018) | 23, (50 %) | Radiographic | Gait retraining (n=12)<br><br>Walking exercise (n=11) | 24.5(2.4)<br>GI 2, GII 8,<br>60.8(6.4)<br><br>25.2(1.1)<br>GI 3, GII 7<br>63.1(5.9) | 6 weeks | Optical (8 camera Vicon system) with AMTI force sensing treadmill | Gait<br>1st PKAM<br>PKFM | WOMAC function<br>WOMAC pain |
| (DeVita et al., 2018) | 16, (67 %) | ACR | Quadriceps strengthening | 26.4(4)<br>GI 2, GII 4, GIII 7, GIV 2,<br>58.1(6.5) | 12 weeks | Optical (Qualisys system) with AMTI force plates | Gait<br>PKEM<br>PKFA early stance | WOMAC function<br>WOMAC pain |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (Fisher et al., 1997) | 10, (100 %) | Radiographic | Progressive resistance training | Ht 1.63 (0.5), Wt 75.5(16.8), >GII (not specified), 62.1(7.9) | 2 months | Video (Peak Performance) with Kistler force plates | Gait<br>PKEM<br>PKFA early stance | Jette Functional Status Index<br>Difficulty<br>Pain |
| (Foroughi et al., 2011a) | 54, (100 %) | Radiographic (Modified Outerbridge Classification) | Progressive resistance training (n=26)<br>Sham exercise (n=28) | 31.4(5.4), GI 9, GII 5, GIII 10, GIV 1, 66(8)<br>32.7(8.4), GI 11, GII 2, GIII 10, GIV 2 65(7) | 6 months | Optical (10 camera Motion Analysis Corporation system) with Kistler force plates | Gait<br>1st PKAM<br>2nd PKAM<br>PKEM<br>PKFM | WOMAC function<br>WOMAC pain |
| (Foroughi et al., 2011) | 54, (100 %) | Radiographic (Modified Outerbridge Classification) | Progressive resistance training<br>Sham exercise | 31.9(5.2), GI 6, GII 5, GIII 5, GIV 10, 64(7)<br>33.2(8.1), GI 12, GII 3, GIII 7, GIV 6, 64(8) | 6 months | Optical (10 camera Motion Analysis Corporation system) with Kistler force plates | Gait<br>PKAddA<br>1st PKAM@ | WOMAC function<br>WOMAC pain |
| (Gaudreault et al., 2011) | 29, (76 %) | ACR | Quadriceps strengthening and proprioception exercises<br>Manual therapy | 31(5), GI 10, GII 5, GIII 5, GIV 9, 63.3(8.4) | 12 weeks | Optical (6 camera Vicon system) with Kistler force sensing treadmill | Gait<br>1st PKAM<br>PKAM Impulse<br>PKFM<br>2nd PKFM<br>PKEM<br>PKIRM<br>PKERM<br>PKFA<br>PKFA early stance<br>Flexion range<br>PKAddA<br>PKAbdA<br>Adduction-Abduction range<br>PKIRA<br>PKERA<br>IR-ER range | WOMAC function<br>WOMAC pain |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (Holsgaard-Larsen et al., 2017) | 47, (62 %) | ACR | Neuromuscular exercise (functional, proprioceptive, endurance strengthening) | 27(3),<br>GI 26, GII 14, GIII 7,<br>57.9(7.9) | 8 weeks | Optical (8 camera Vicon system) with AMTI force plates | Gait<br>1st PKAM<br>KAM Impulse | KOOS ADL<br>KOOS pain |
| (Hunt et al., 2018) | 79, (70 %) | Radiography | Gait retraining (toe-out) (n=40) | 27.3(3.5),<br>GII 19, GIII 17, GIV 4,<br>64.6(7.6) | 4 months | Optical (10 camera Motion Analysis Corporation system) with AMTI force plates | Gait<br>1st PKAM<br>2nd PKAM<br>KAM Impulse<br>PKFM | WOMAC function<br>WOMAC pain<br>NRS past month |
| | | | Progressive walking (n=39) | 27.4(3.5),<br>GII 18, GIII 14, GIV 7,<br>65.4(9.6 | | | | |
| (Hunt & Takacs, 2014) | 16, (56 %) | Radiography | Gait retraining (toe-out) | 29.9(6.8),<br>GII 4, GIII 9, GIV 3,<br>64.8(10.4) | 10 weeks | Optical (10 camera Motion Analysis Corporation system) with AMTI force plates | Gait<br>1st PKAM<br>2nd PKAM<br>KAM impulse<br>PKFM | WOMAC function<br>WOMAC pain<br>NRS past month |
| (King et al., 2008) | 14, (14 %) | ACR | Resistance training (seated) | 29.3(3.3),<br>GI 2, GII 4, GIII 7, GIV 1,<br>48.4(6.5) | 12 weeks | Optical (8 camera Motion Analysis Corporation system) with AMTI force plates | Gait<br>1st PKAM | KOOS ADL<br>KOOS Pain<br>NRS Tibiofemoral pain<br>ADL |
| (Lim et al., 2008) | 53, (57 %) | ACR | Quadriceps resistance exercise (sitting and supine) | Neutral group<br>29(5.2),<br>GII 12, GIII 7, GIV 8,<br>64.1(9.3) | 12 weeks | Optical (8 camera Vicon system) with AMTI force plates | Gait<br>1st PKAM | WOMAC function<br>WOMAC pain |
| | | | | Malaligned group<br>28.2(3.7),<br>GII 4, GIII 8, GIV 14,<br>67.2(6.7) | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (Preece et al., 2016) | 22, (50 %) | Radiography | Alexander technique | 29(4),<br>GII 11, GIII 3, GIV 4,<br>62(10) | 12 weeks | Optical (10 camera Qualisys system) with AMTI force plates<br>Surface EMG (Telemyo system) | Gait<br>EMG Co-contraction<br>Pre-contact<br>Early stance<br>Lateral - VL, BF<br>Medial – VM, ST | WOMAC Pain |
| (Roper et al., 2013) | 14, (86 %) | ACR | Aquatic treadmill<br>Land treadmill | 33.5(8.4),<br>NR,<br>59.2(7.2) | 1 week | Optical (7 camera Vicon system) | Gait<br>Maximum angular velocity – Sagittal plane<br>Stance phase<br>Swing phase<br>Minimum angular velocity – Transverse plane<br>Swing phase | VAS pain past week |
| (Shen et al., 2008) | 48, (88 %) | ACR | Tai Chi | 28.1(6.04),<br>NR,<br>64.4(8.3) | 6 weeks | 2D video (Peak Motus) | Gait<br>Flexion range | WOMAC function<br>WOMAC pain<br>VAS pain maximum<br>VAS pain overall |
| (Shull, Silder, et al., 2013) | 10, (40 %) | Radiographic | Gait retraining (toe-in and trunk sway) | 26.6(4.7),<br>GII 3, GIII 6, GIV 1,<br>60(13) | 6 weeks | Optical (8 camera Vicon system) with Bertec force sensing treadmill | Gait<br>1st PKAM<br>PKFM | WOMAC function<br>WOMAC pain<br>VAS pain |
| (Sled et al., 2010) | 40, (58 %) | ACR | Hip strengthening exercise | 27.38(5.47)<br>Mean grade 2.5(9.1)<br>62.98(9.73) | 8 weeks | Optical (2 camera Optotrak system) with AMTI force plates | Gait (speed matched)<br>1st PKAM | WOMAC function<br>WOMAC pain |
| (Turcot et al., 2009) | 24, (75 %) | ACR | Aerobic, proprioceptive/balance and strengthening exercise.<br>Manual therapy, ultrasound, tape | 30.3(4.7),<br>NR,<br>64.2(7.7) | 12 weeks | Optical (6 camera Vicon system) with accelerometers (Physilog) | Gait<br>Acceleration<br>Anterior-posterior<br>Medial-lateral | WOMAC pain |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (Wang et al., 2016) | 50, (53 %) | ACR | Quadriceps strengthening exercise (n=20)<br><br>Quadriceps strengthening exercise + whole body vibration (n=19) | 26.2(2.7)<br>GII 9, GIII 11,<br>61.5(7.3)<br><br>27.8(3.1)<br>GII 9, GIII 10,<br>61.1(7.1) | 12 weeks | Optical (8 camera Vicon system) with AMTI force plates | Gait<br>1st PKAM<br>Peak Valgus Moment<br>KAddA<br>Knee valgus angle | WOMAC function<br>WOMAC pain<br>VAS pain related to joint movement |
| (Zhu et al., 2016) | 23, (100 %) | ACR | Tai Chi | 25.23(3.46),<br>GI 7, GII 12, GIII 4<br>64.61(3.4) | 24 weeks | Optical (16 camera Vicon system) | Gait<br>PKFA early stance<br>Flexion angle at initial contact | WOMAC function<br>WOMAC pain |

Abbreviations: ACR, American College of Rheumatology diagnostic criteria (Altman et al., 1986); BF, biceps femoris; EMG, surface electromyography; ER, external rotation; GI-GIV: grade according to Kellgren Lawrence; IR, internal rotation; KOOS, Knee Injury and Osteoarthritis Outcome Scale; MVIC, maximal voluntary isometric contraction; NR, not reported; NRS, Numerical Rating Scale; PKAbdA, peak knee abduction angle; PKAddA, peak knee adduction angle; PKAM, peak knee adduction moment; PKEM, peak knee extension moment; PKERA, peak knee external rotation angle; PKERM, peak knee external rotation moment; PKFA, peak knee flexion angle; PKFM, peak knee flexion moment; PKIRA, peak knee internal rotation angle; PKIRM, peak knee internal rotation moment; ST, semitendinosus; VAS, Visual Analogue Scale; VL, vastus lateralis; VM, vastus medialis; WOMAC; Western Ontario McMaster Universities Osteoarthritis Index. @excluded from review as reported in (Foroughi et al., 2011a).

## Appendix 3-5. Pre, post and change scores of all investigated gait moment parameters.

| Author & year (Cohort or alternate measure) | n | Time between measures | Movement Parameter | | | |
|---|---|---|---|---|---|---|
| | | | Pre Mean (SD) | Post Mean (SD) | Change Mean (SD) 95%CI (or p value) | Standardised response mean |
| Moments (Nm/kg.ht%)# | | | | | | |
| Changed - 1st Peak Knee Adduction Moment | | | | | | |
| Bennell 2010 (Neuromuscular exercise) | 45 | 13 weeks | 3.2 (1.00) | 3.3 (0.90) | 0.15 (0.44) 0.01 to 0.29 | 0.34 |
| Cheung 2018 (Gait retraining – toe out) | 12 | 6 weeks | 0.35 (0.05) | 0.27 (0.05) | -0.08 (0.03) -0.10 to -0.05 | -2.62 |
| Holsgaard-Larsen 2017 (Neuromuscular exercise) | 47 | 8 weeks | 2.86 (0.83) | 2.98 (Not reported) | 0.12 (0.31) 0.03 to 0.22 | 0.38 |
| Hunt 2018 (Progressive walking) | 39 | 4 months | 2.38 (0.88) | 2.57 (0.37) | 0.13 (0.37) 0.00 to 0.26 | 0.35 |
| Lim 2008 (Quads strength – more neutrally aligned) | 27 | 12 weeks | 3.58 (0.94) | 3.63 (1.11) | 0.05 (0.0 9) 0.01 to 0.09 | 0.56 |
| Lim 2008 (Quads strength – more malaligned) | 26 | 12 weeks | 4.28 (0.63) | 4.4 (0.76) | 0.12 (0.09) 0.08 to 0.16 | 1.33 |
| Shull 2013 (Gait retraining – toe in) | 10 | 6 weeks | 3.11 (1.40) | 2.61 (1.47) | -0.50 (0.48) -0.84 to -0.16 | -0.34 |
| No change - 1st Peak Knee Adduction Moment | | | | | | |
| Bennell 2014 (Neuromuscular exercise) | 50 | 13 weeks | 3.05 (0.90) | 3.26 (0.95) | 0.12 (0.50) -0.05 to 0.29 | 0.24 |
| Bennell 2014 (Quads strengthening) | 50 | 13 weeks | 3.21 (0.88) | 3.30 (0.79) | -0.04 (0.46) -0.18 to 0.10 | -0.09 |
| Brenneman 2015 (Yoga)^ | 45 | 12 weeks | 0.42 (0.16) | 0.43 (0.15) | 0.01 (0.09) -0.02 to 0.04 | 0.11 |
| Cheung 2018 (Progressive walking) | 11 | 6 weeks | 0.32 (0.06) | 0.33 (0.05) | 0.01 (0.02) -0.01 to 0.02 | 0.14 |
| Foroughi 2011a (Strengthening) | 26 | 6 months | -2.63 (1.26) | -2.65 (1.26) | -0.02 (0.42) -0.21 to 0.17 | -0.02 |
| Foroughi 2011a (Sham exercise) | 28 | 6 months | -2.43 (1.06) | -2.54 (0.98) | -0.11 (0.35) -0.25 to 0.03 | -0.1 |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | -2.35 (1.19) | -2.45 (1.03) | -0.10 (0.40) -0.25 to 0.05 | -0.08 |
| Hunt 2014 (Gait retraining – toe out) | 16 | 10 weeks | 3.45 (0.82) | 3.19 (0.72) | -0.26 (0.60) -0.60 to 0.07 | -0.43 |
| Hunt 2018 (Gait retraining – toe out) | 40 | 4 months | 2.41 (1.30) | 2.43 (0.37) | -0.01 (0.36) -0.13 to 0.11 | -0.03 |
| King 2008 (Resistance training) | 14 | 12 weeks | 3.3 (0.72) | 3.43 (0.49) | 0.13 (0.47) -0.14 to 0.4 | 0.28 |
| Sled 2010 (Hip strengthening) | 40 | 8 weeks | 2.97 (0.84) | 2.96 (0.87) | -0.01 (0.29) -0.10 to 0.08 | -0.01 |
| Wang 2016^ (Vibration + Quads strengthening) | 19 | 12 weeks | 0.53 (0.20) | 0.52 (0.15) | -0.01 (0.08) -0.05 to 0.03 | -0.04 |
| Wang 2016^ (Quads strengthening) | 20 | 12 weeks | 0.54 (0.14) | 0.53 (0.17) | -0.01 (0.06) -0.04 to 0.02 | -0.06 |
| No change - Peak Knee Flexion Moment | | | | | | |
| Bennell 2014 (Neuromuscular exercise) | 50 | 13 weeks | 4.02 (1.38) | 3.89 (1.64) | -0.03 (1.08) -0.38 to 0.32 | -0.03 |
| Bennell 2014 (Quads strengthening) | 50 | 13 weeks | 3.96 (1.59) | 4.05 (1.79) | 0.07 (0.82) -0.18 to 0.32 | 0.08 |
| Cheung 2018 (Gait retraining) | 12 | 6 weeks | 0.30 (0.04) | 0.29 (0.04) | -0.01 (0.03) -0.03 to 0.01 | -0.16 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Cheung 2018 (Progressive walking) | 11 | 6 weeks | 0.27 (0.05) | 0.27 (0.05) | -0.00 (0.04) -0.03 to 0.02 | -0.02 |
| Foroughi 2011a (Strengthening) | 26 | 6 months | 3.56 (0.95) | 3.62 (0.97) | 0.06 (0.78) -0.26 to 0.38 | -0.06 |
| Foroughi 2011a (Sham exercise) | 28 | 6 months | 3.26 (1.72) | 3.32 (1.66) | 0.06 (1.37) -0.44 to 0.56 | 0.03 |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | -1.96 (1.27) | -2.2 (1.34) | -0.24 (1.06) -0.59 to 0.11 | -0.18 |
| Hunt 2014 (Gait retraining – toe out) | 16 | 10 weeks | 1.38 (1.36) | 1.51 (1.29) | 0.13 (1.18) -0.52 to 0.78 | 0.11 |
| Hunt 2018 (Gait retraining – toe out) | 40 | 4 months | 3.01 (1.40) | 3.14 (0.97) | 0.06 (0.96) -0.26 to 0.38 | 0.06 |
| Hunt 2018 (Progressive walking) | 39 | 4 months | 3.2 (1.52) | 3.18 (0.98) | 0.09 (0.94) -0.24 to 0.42 | 0.09 |
| Shull 2013 (Gait retraining – toe in) | 10 | 6 weeks | 1.95 (0.76) | 1.67 (0.75) | -0.28 (0.61) -0.66 to 0.10 | -0.37 |
| Gaudreault 2011% (Exercise therapy) | 29 | 12 weeks | -1.1 (0.60) | -1.2 (0.60) | -0.1 (0.49) -0.28 to 0.08 | -0.17 |
| Changed - Peak Knee Adduction Moment Impulse& | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 94.1 (11.3) | 88.2 (11.7) | -5.90 (6.78) -8.48 to -3.32 | -0.87 |
| Holsgaard-Larsen 2017 (Neuromuscular exercise) | 47 | 8 weeks | 1.09 (0.45) | Not reported | 0.05 (0.13) 0.01 to 0.09 | 0.38 |
| Hunt 2018 (Gait retraining – toe out) | 40 | 4 months | 0.84 (0.11) | 0.82 (0.12) | -0.04 (0.10) -0.07 to -0.01 | -0.38 |
| No change - Peak Knee Adduction Moment Impulse& | | | | | | |
| Bennell 2010 (Neuromuscular exercise) | 45 | 13 weeks | 1.10 (0.40) | 1.10 (0.40) | 0.05 (0.19) -0.01 to 0.11 | 0.27 |
| Bennell 2014 (Neuromuscular exercise) | 50 | 13 weeks | 1.15 (0.37) | 1.20 (0.36) | 0.02 (0.21) -0.05 to 0.09 | 0.09 |
| Bennell 2014 (Quads strengthening) | 50 | 13 weeks | 1.21 (0.36) | 1.23 (0.37) | -0.02 (0.18) -0.07 to 0.03 | -0.11 |
| Hunt 2014 (Gait retraining – toe out) | 16 | 10 weeks | 1.33 (0.29) | 1.24 (0.34) | -0.08 (0.22) -0.20 to 0.04 | -0.37 |
| Hunt 2018 (Progressive walking) | 39 | 4 months | 0.86 (0.38) | 0.87 (0.44) | 0.01 (0.10) -0.02 to 0.04 | 0.1 |
| Changed – Peak Knee Extension Moment | | | | | | |
| DeVita 2018^ (Strengthening) | 16 | 12 weeks | 0.52 (0.27) | 0.64 (0.28) | 0.12 (0.05) 0.04 to 0.19 | 0.43 |
| Foroughi 2011a (Sham exercise) | 28 | 6 months | -1.74 (0.52) | -2.21 (0.60) | -0.47 (0.30) -0.59 to -0.35 | -0.81 |
| No change – Peak Knee Extension Moment | | | | | | |
| Fisher 1997^ (Strengthening) | 10 | 2 months | 0.26 (0.10) | 0.24 (0.09) | -0.02 (0.05) -0.06 to 0.02 | -0.21 |
| Foroughi 2011a (Strengthening) | 26 | 6 months | -1.81 (0.44) | -1.88 (0.41) | -0.07 (0.22) -0.17 to 0.03 | -0.16 |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 0.66 (1.50) | 0.49 (1.35) | -0.17 (0.74) -0.45 to 0.11 | -0.12 |
| Changed - 2nd Peak Knee Adduction Moment | | | | | | |
| Hunt 2018 (Gait retraining – toe out) | 40 | 4 months | 2.67 (0.40) | 2.44 (0.30) | -0.24 (0.27) -0.33 to -0.15 | -0.89 |
| Hunt (2014) (Gait retraining – toe out) | 16 | 10 weeks | 2.87 (0.92) | 2.57 (0.84) | -0.3 (0.49) -0.57 to -0.03 | -0.62 |
| No change - 2nd Peak Knee Adduction Moment | | | | | | |
| Foroughi 2011a (Strengthening) | 26 | 6 months | -2.14 (1.34) | -2.00 (1.28) | 0.14 (1.02) -0.34 to 0.62 | 0.07 |
| Foroughi 2011a (Sham exercise) | 28 | 6 months | -1.84 (0.94) | -1.78 (0.91) | 0.06 (0.89) -0.31 to 0.43 | 0.03 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Hunt 2018 (Progressive walking) | 39 | 4 months | 2.63 (1.01) | 2.69 (1.19) | 0.02 (0.27) -0.07 to 0.11 | 0.07 |
| No change - Peak Knee Internal Rotation Moment | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 0.66 (0.28) | 0.61 (0.30) | -0.05 (0.18) -0.12 to 0.02 | 0.11 |
| Changed – Peak Knee Abduction Moment | | | | | | |
| Wang 2016^ (Vibration + Quads strengthening) | 19 | 12 weeks | 0.11 (0.04) | 0.09 (0.06) | -0.02 (0.02) -0.03 to -0.01 | -0.25 |
| No Change – Peak Knee Abduction Moment | | | | | | |
| Wang 2016^ (Quads strengthening) | 20 | 12 weeks | 0.11 (0.05) | 0.09 (0.07) | -0.02 (0.09) -0.06 to 0.02 | -0.07 |
| No change - Peak Knee External Rotation Moment | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | -0.27 (0.28) | -0.25 (0.32) | 0.02 (0.18) -0.05 to 0.09 | 0.11 |
| Kinematics (°)# | | | | | | |
| No change - Knee Flexion Range | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 58.0 (8.5) | 59.5 (8.3) | 1.50 (4.63) -0.26 to 3.26 | 0.18 |
| Shen 2008 (Tai Chi) | 48 | 6 weeks | 54.8 (6.6) | 55.6 (5.7) | 0.80 (3.5) -0.32 to 1.92 | 0.13 |
| Changed - Peak Knee Flexion during Early Stance | | | | | | |
| Zhu 2016 (Tai chi) | 23 | 24 weeks | 18.01 (3.65) | 21.03 (2.78) | 3.02 (2.39) 1.93 to 4.11 | 1.26 |
| No change - Peak Knee Flexion during Early Stance | | | | | | |
| DeVita 2018 (Strengthening) | 16 | 12 weeks | -16.4 (7.50) | -16.4 (5.60) | 0 (4.23) -2.34 to 2.34 | 0 |
| Fisher 1997 (Strengthening) | 10 | 2 months | 11.5 (6.60) | 10.4 (5.60) | -1.10 (3.68) -4.51 to 2.31 | -0.17 |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 18.10 (6.80) | 18.60 (7.30) | 0.50 (4.14) -1.07 to 2.07 | 0.07 |
| No change - Peak Knee Adduction Angle | | | | | | |
| Foroughi 2011b (Strengthening) | 26 | 6 months | 1.87 (6.90) | 0.14 (7.58) | -0.82 (2.27) -1.95 to 0.31 | -0.23 |
| Foroughi 2011b (Sham exercise) | 28 | 6 months | 3.49 (5.32) | 2.67 (4.78) | -0.5 (2.18) -1.55 to 0.55 | -0.16 |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | 6.50 (5.00) | 6.00 (5.00) | -0.32 (1.25) -0.79 to 0.15 | -0.1 |
| Wang 2016 (Vibration + Quads strengthening) | 19 | 12 weeks | 8.80 (3.23) | 8.75 (3.12) | -0.05 (1.39) -0.72 to 0.62 | -0.02 |
| Wang 2016 (Quads strengthening) | 20 | 12 weeks | 8.85 (3.54) | 8.79 (3.33) | -0.06 (1.51) -0.77 to 0.65 | -0.02 |
| Changed - Peak Knee Abduction Angle | | | | | | |
| Wang 2016 (Vibration + Quads strength) | 19 | 12 weeks | 5.14 (0.50) | 4.89 (0.60) | -0.25 (0.23) -0.36 to -0.14 | -0.41 |
| Wang 2016 (Quads strength) | 20 | 12 weeks | 5.2 (0.43) | 4.9 (0.43) | -0.3 (0.21) -0.4 to -0.2 | -0.53 |
| No change - Peak Knee Abduction Angle | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 29 | 12 weeks | -2.3 (5.60) | -3.1 (6.60) | -0.8 (2.52) -1.76 to 0.16 | -0.12 |
| Changed - Maximum Sagittal Angular Velocity during Stance^ | | | | | | |
| Roper 2013 (Aquatic treadmill) | 14 | 1 week | 207 (47.30) | 251 (36.20) | 44.3 (57.4) 11.1 to 77.4 | 0.77 |
| No change - Maximum Sagittal Angular Velocity during Stance^ | | | | | | |
| Roper 2013 (Treadmill) | 14 | 1 week | 226 (88.70) | 188 (51.40) | -23.7 (58.8) -57.6 to 10.25 | -0.4 |
| Changed - Minimum Transverse Angular Velocity^ | | | | | | |
| Roper 2013 (Aquatic treadmill) | 14 | 1 week | 201 (95.40) | 293 (109.0) | 91.4 (93.9) 37.2 to 146 | 0.5 |
| No change - Minimum Transverse Angular Velocity^ | | | | | | |
| Roper 2013 (Treadmill) | 14 | 1 week | 224 (107) | 181 (88.8) | -27.6 (56.2) -60 to 4.85 | -0.54 |
| No change - Max Sagittal Angular Velocity Swing^ | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Roper (Aquatic treadmill) | 2013 | 14 | 1 week | 299 (72.1) | 337 (33.7) | 38.1 (76.7) -6.18 to 82.38 | 0.5 |
| Roper (Treadmill) | 2013 | 14 | 1 week | 315 (71.1) | 292 (74) | -25.5 (47.1) -52.7 to 1.69 | -0.54 |
| No change - Peak Knee External Rotation Angle | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | | 29 | 12 weeks | 5.2 (4.10) | 4.9 (3.90) | -0.3 (4.21) -1.90 to 1.30 | -0.07 |
| No change - Peak Flexion Angle during Swing | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | | 29 | 12 weeks | 65.1 (4.5) | 66 (4.7) | 0.90 (2.54) -0.07 to 1.87 | 0.19 |
| No change - Peak Knee Internal Rotation Angle | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | | 29 | 12 weeks | -3.2 (2.90) | -3.7 (2.70) | -0.5 (2.95) -1.62 to 0.62 | -0.18 |
| Changed - Acceleration Anterior-Posterior& | | | | | | | |
| Turcot 2009 (Exercise therapy) | | 24 | 12 weeks | -0.88 (0.42) | -0.74 (0.38) | 0.14 p=0.02 | 0.52 |
| Changed - Knee Angle at Initial Contact | | | | | | | |
| Zhu 2016 (Tai chi) | | 23 | 24 weeks | 2.56 (3.10) | 5.23 (2.50) | 2.67 (3) 1.30 to 4.03 | 0.89 |
| No change - Acceleration Medial-Lateral& | | | | | | | |
| Turcot 2009 (Exercise therapy) | | 24 | 12 weeks | -0.56 (0.33) | -0.57 (0.30) | -0.01 p=0.86 | 0.05 |
| No change– Adduction-abduction range | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | | 29 | 12 weeks | 8.8 (3.60) | 9.0 (3.40) | 0.2 (1.34) -0.31 to 0.71 | 0.06 |
| No change - Internal-external rotation range | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | | 29 | 12 weeks | 8.3 (2.80) | 8.5 (3.30) | 0.2 (3.23) -1.03 to 1.43 | 0.06 |

Change was defined as 95% confidence interval does not cross zero and standardised response mean >0.2.
Abbreviation: CI, confidence interval; KFM, knee flexion moment PKAM; SD, standard deviation. #Movement parameter reported in ° unless otherwise indicated (^°/second, &g).

## Appendix 3-6. Knee joint moment parameter standardised response mean compared with activity limitation and pain change for movement parameters that demonstrated a change in at least one study.

| Author & year | Movement parameter | Activity Limitation (improvement) | | | | | Pain (improvement) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | SRM | Outcome measure | Pre Mean (SD) | Post Mean (SD) | Change Mean (SD) 95%CI or p value | SRM | Outcome measure | Pre Mean (SD) | Post Mean (SD) | Change Mean (SD) 95%CI or p value | SRM |
| Changed - 1st Peak Knee Adduction Moment | | | | | | | | | | | |
| Bennell 2010 (Neuromuscular exercise) | 0.34 | WOMAC function | 24.8 (10.9) | 16.2 (11.2) | 8.6 (7.18) 6.27 to 10.93 | 1.2 | WOMAC pain | 7.7 (3) | 4.9 (3.3) | 2.8 (2.44) 2.01 to 3.59 | 1.15 |
| | | | | | | | NRS pain on walking | 4.3 (2) | 2.6 (2.1) | 1.7 (1.75) 1.13 to 2.27 | 0.97 |
| Cheung 2018 (Gait retraining) | -2.62 | WOMAC function | 19.9 (4.9) | 11.4 (5.4) | 8.57 (3.01) 6.41 to 10.72 | 2.84 | WOMAC pain | 6 (2.5) | 3.06 (2) | 2.94 (1.34) 1.98 to 3.9 | 2.19 |
| Holsgaard-Larsen 2017 (Neuromuscular exercise) | 0.38 | KOOS ADL | 68.2 (15.5) | Not reported | 6.96 (10) 3.76 to 10.16 | 0.69 | KOOS pain | 61.6 (13.7) | Not reported | 7.23 (9.63) 4.15 to 10.31 | 0.75 |
| Hunt 2018 (Progressive walking) | 0.35 | WOMAC function | 21.4 (9.49) | 16.7 (9.47) | 7.7 (9.46) 4.5 to 10.9 | 0.81 | WOMAC pain | 6.4 (2.5) | 5.4 (3) | 1.5 (2.95) 0.5 to 2.5 | 0.51 |
| | | | | | | | NRS average pain past week | 3.7 (1.87) | 2.3 (1.8) | 1.8 (1.77) 1.2 to 2.4 | 1.01 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lim 2008 (more neutrally aligned) | 0.56 | WOMAC function | 22.5 (10.5) | 16.3 (12.3) | 6.26 (6.99) 3.30 to 9.21 | 0.89 | WOMAC pain | 7.14 (2.92) | 4.56 (3.38) | 2.6 (2.25) 1.65 to 3.55 | 1.15 |
| Lim 2008 (more malaligned) | 1.33 | WOMAC function | 21.4 (11.8) | 19.9 (10.6) | 1.43 (7.14) -1.5 to 4.37 | No change | WOMAC pain | 6.62 (3.08) | 5.7 (3.38) | 0.92 (2.5) -0.1 to 1.95 | No change |
| Shull 2013 (Gait retraining – toe in) | -0.34 | WOMAC function | 15.6 (15) | 6.1 (4.8) | 9.52 (11.94) 0.97 to 18.06 | 0.59 | WOMAC pain | 6 (3.6) | 3 (2.2) | 3 (2.70) 1.07 to 4.93 | 0.91 |
| | | | | | | | NRS (worst pain) | 3.2 (2.3) | 1.3 (0.9) | 1.9 (1.93) 0.52 to 3.28 | 0.91 |
| No change - 1st Peak Knee Adduction Moment | | | | | | | | | | | |
| Bennell 2014 (Neuromuscular exercise) | 0.24 | WOMAC function | 26 (9.1) | 18.3 (9.6) | 7.5 (7.76) 6.27 to 10.93 | 0.97 | WOMAC pain | 8.1 (2.2) | 6.4 (3.1) | 2.8 (2.43) 2.01 to 3.59 | 1.15 |
| | | | | | | | VAS overall pain | 54 (13.3) | 34.1 (23.6) | 19.9 (21.3) 13 to 26.8 | 0.97 |
| | | | | | | | VAS pain on walking | 59.5 (15) | 39.6 (25.9) | 19.9 (23.88) 12.05 to 27.75 | 0.83 |
| Bennell 2014 (Quads strengthening) | -0.09 | WOMAC function | 28.2 (9.9) | 20.1 (9.8) | 8.1 (7.89) 5.7 to 10.5 | 0.92 | WOMAC pain | 8.8 (3.3) | 6.4 (2.9) | 1.7 (2.59) 0.85 to 2.55 | 0.66 |
| | | | | | | | VAS overall pain | 54.2 (16.8) | 31.4 (19.3) | 22.8 (19.4) 16.42 to 29.18 | 1.17 |
| | | | | | | | VAS pain on walking | 55.3 (22.4) | 40 (22.9) | 15.3 (22.69) 8.4 to 22.2 | 0.67 |
| Brenneman 2015 (Yoga) | 0.11 | KOOS ADL | 74.9 (15.8) | 87.1 (11.1) | 11.3 (14.65) 6.48 to 16.12 | 0.77 | KOOS pain | 67.7 (15.4) | 79.4 (12.7) | 12 (12.65) 7.84 to 16.16 | 0.95 |
| Cheung 2018 (Progressive walking) | 0.14 | WOMAC function | 17 (4.96) | 15.3 (4.76) | 1.7 (3.58) -0.86 to 4.26 | No change | WOMAC pain | 4.4 (2.2) | 4 (2.2) | 0.4 (1.81) -0.89 to 1.69 | No change |
| Foroughi 2011a (Strengthening) | -0.02 | WOMAC function | 19.4 (9.8) | 13.3 (9.4) | 6.1 (7.08) 2.78 to 9.41 | 0.63 | WOMAC pain | 5.6 (3.2) | 3.8 (2.7) | 1.8 (2.46) 0.65 to 2.95 | 0.6 |
| Foroughi 2011a (Sham exercise) | -0.1 | WOMAC function | 23.3 (11.3) | 18.1 (12) | 5.2 (8.61) 1.65 to 8.75 | 0.44 | WOMAC pain | 5.6 (3.2) | 4.4 (3.7) | 1.2 (2.87) 0.02 to 2.38 | 0.34 |
| Gaudreault 2011 (Exercise therapy) | -0.08 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Hunt 2014 (Gait retraining – toe out) | -0.43 | NA | - | - | - | - | WOMAC pain | 7.4 (3.4) | 5.3 (2.9) | 2.1 (3.16) 0.28 to 3.92 | 0.66 |
| | | | | | | | NRS pain over previous week | 4.5 (1.7) | 2.6 (1.8) | 1.9 (1.62) 0.96 to 2.84 | 1.17 |
| Hunt 2018 (Gait retraining – toe out) | -0.03 | WOMAC function | 28.1 (11.9) | 13 (9.6) | 11.4 (9.58) 8.25 to 14.55 | 1.19 | WOMAC pain | 7.6 (3.16) | 4.2 (2.96) | 2.7 (2.89) 1.75 to 3.65 | 0.93 |
| | | | | | | | NRS average pain past week | 4.7 (2.53) | 2 (1.77) | 2.1 (1.82) 1.5 to 2.7 | 1.15 |
| King 2008 (Resistance training) | 0.28 | KOOS ADL | 72.47 (16.54 ) | 73.21 (14.41 ) | 0.74 (9.76) -4.9 to 6.37 | No change | KOOS pain | 59.52 (17.25 ) | 62.28 (15.78 ) | 2.76 (7.87) -1.79 to 7.31 | No change |
| | | | | | | | NRS ADL pain | 1.76 (1.82) | 1.69 (1.79) | 0.07 (1.45) -0.77 to 0.91 | No change |
| Sled 2010 (Hip strengthening) | -0.01 | WOMAC function | 19.6 (11.4) | 18.15 (12.4) | 1.45 (8.81) -1.37 to 4.27 | No change | WOMAC pain | 5.55 (2.78) | 4.78 (3.31) | 0.77 (2.55) -0.05 to 1.59 | No change |
| Wang 2016 (Vibration + Quads strengthening) | -0.04 | WOMAC function | 35.7 (10.3) | 12.5 (9.7) | 23.2 (7.39) 19.64 to 26.76 | 2.31 | WOMAC pain | 13.2 (3.5) | 6.7 (2.8) | 6.5 (2.66) 5.22 to 7.78 | 2 |
| | | | | | | | VAS pain on joint movement | 8.2 (1.1) | 3.4 (0.9) | 4.8 (0.86) 4.38 to 5.22 | 4.69 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Wang 2016 (Quads strengthening) | -0.06 | WOMAC function | 33.8 (11.2) | 15 (8.2) | 18.8 (7.67) 15.21 to 22.39 | 1.8 | WOMAC pain | 12.7 (4) | 1.9 (0.5) | 10.8 (3.69) 9.07 to 12.53 | 2.4 |
| | | | | | | | VAS pain on joint movement | 8 (1.2) | 3.8 (0.9) | 4.2 (0.93) 3.77 to 4.63 | 3.82 |
| Changed - Peak Knee Adduction Moment Impulse | | | | | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | -0.87 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Holsgaard-Larsen 2017 (Neuromuscular exercise) | 0.38 | KOOS ADL | 68.2 (15.5) | Not reported | 6.96 (10) 3.76 to 10.16 | 0.69 | KOOS pain | 61.6 (13.7) | Not reported | 7.23 (9.63) 4.15 to 10.31 | 0.75 |
| Hunt 2018 (Gait retraining – toe out) | -0.38 | WOMAC function | 28.1 (11.9) | 13 (9.6) | 11.4 (9.58) 8.25 to 14.55 | 1.19 | WOMAC pain | 7.6 (3.16) | 4.2 (2.96) | 2.7 (2.89) 1.75 to 3.65 | 0.93 |
| | | | | | | | NRS average pain past week | 4.7 (2.53) | 2 (1.77) | 2.1 (1.82) 1.5 to 2.7 | 1.15 |
| No change - Peak Knee Adduction Moment Impulse | | | | | | | | | | | |
| Bennell 2010 (Neuromuscular exercise) | 0.27 | WOMAC function | 24.8 (10.9) | 16.2 (11.2) | 8.6 (7.18) 6.27 to 10.93 | 1.2 | WOMAC pain | 7.7 (3) | 4.9 (3.3) | 2.8 (2.44) 2.01 to 3.59 | 1.15 |
| | | | | | | | NRS pain on walking | 4.3 (2) | 2.6 (2.1) | 1.7 (1.75) 1.13 to 2.27 | 0.97 |
| Bennell 2014 (Neuromuscular exercise) | 0.09 | WOMAC function | 26 (9.1) | 18.3 (9.6) | 7.5 (7.76) 6.27 to 10.93 | 0.97 | WOMAC pain | 8.1 (2.2) | 6.4 (3.1) | 2.8 (2.43) 2.01 to 3.59 | 1.15 |
| | | | | | | | VAS overall pain | 54 (13.3) | 34.1 (23.6) | 19.9 (21.3) 13 to 26.8 | 0.97 |
| | | | | | | | VAS pain on walking | 59.5 (15) | 39.6 (25.9) | 19.9 (23.88) 12.05 to 27.75 | 0.83 |
| Bennell 2014 (Quads strengthening) | -0.11 | WOMAC function | 28.2 (9.9) | 20.1 (9.8) | 8.1 (7.89) 5.7 to 10.5 | 0.92 | WOMAC pain | 8.8 (3.3) | 6.4 (2.9) | 1.7 (2.59) 0.85 to 2.55 | 0.66 |
| | | | | | | | VAS overall pain | 54.2 (16.8) | 31.4 (19.3) | 22.8 (19.4) 16.42 to 29.18 | 1.17 |
| | | | | | | | VAS pain on walking | 55.3 (22.4) | 40 (22.9) | 15.3 (22.69) 8.4 to 22.2 | 0.67 |
| Hunt 2014 (Gait retraining – toe out) | -0.37 | NA | - | - | - | - | WOMAC pain | 7.4 (3.4) | 5.3 (2.9) | 2.1 (3.16) 0.28 to 3.92 | 0.66 |
| | | | | | | | NRS pain over previous week | 4.5 (1.7) | 2.6 (1.8) | 1.9 (1.62) 0.96 to 2.84 | 1.17 |
| Hunt 2018 (Progressive walking) | 0.1 | WOMAC function | 21.4 (9.49) | 16.7 (9.47) | 7.7 (9.46) 4.5 to 10.9 | 0.81 | WOMAC pain | 6.4 (2.5) | 5.4 (3) | 1.5 (2.95) 0.5 to 2.5 | 0.51 |
| | | | | | | | NRS average pain past week | 3.7 (1.87) | 2.3 (1.8) | 1.8 (1.77) 1.2 to 2.4 | 1.01 |
| Changed - Peak Knee Extension Moment | | | | | | | | | | | |
| DeVita 2018 (Strengthening) | 0.43 | WOMAC function | 17.7 (9.7) | 6.8 (9.5) | 10.9 (7.07) 6.32 to 15.14 | 1.13 | WOMAC pain | 6.43 (3.18) | 3.64 (2.55) | 2.79 (2.42) 1.45 to 4.13 | 0.95 |
| Foroughi 2011a (Sham exercise) | -0.81 | WOMAC function | 23.3 (11.3) | 18.1 (12) | 5.2 (8.61) 1.65 to 8.75 | 0.44 | WOMAC pain | 5.6 (3.2) | 4.4 (3.7) | 1.2 (2.87) 0.02 to 2.38 | 0.34 |
| No change – Peak Knee Extension Moment | | | | | | | | | | | |
| Fisher 1997 (Strengthening)^ | -0.21 | Jette Functional Index - Difficulty | 2.2 (0.8) | 1.9 (0.7) | 0.29 (0.59) -0.84 to 0.25 | No change | Jette Functional Index - Pain | 1.79 (0.53) | 1.40 (0.25) | 0.38 (0.41) 0.76 to 0.01 | 0.78 |
| Foroughi 2011a (Strengthening) | -0.16 | WOMAC function | 19.4 (9.8) | 13.3 (9.4) | 6.1 (7.08) 2.78 to 9.41 | 0.63 | WOMAC pain | 5.6 (3.2) | 3.8 (2.7) | 1.8 (2.46) 0.65 to 2.95 | 0.6 |
| Gaudreault 2011 (Exercise therapy) | -0.12 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Changed - 2nd Peak Knee Adduction Moment | | | | | | | | | | | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Hunt 2014 (Gait retraining – toe out) | -0.89 | NA | - | - | - | - | WOMAC pain | 7.4 (3.4) | 5.3 (2.9) | 2.1 (3.16) 0.28 to 3.92 | 0.66 |
| | | | | | | | NRS pain over previous week | 4.5 (1.7) | 2.6 (1.8) | 1.9 (1.62) 0.96 to 2.84 | 1.17 |
| Hunt 2018 (Gait retraining – toe out) | -0.62 | WOMAC function | 28.1 (11.9) | 13 (9.6) | 11.4 (9.58) 8.25 to 14.55 | 1.19 | WOMAC pain | 7.6 (3.16) | 4.2 (2.96) | 2.7 (2.89) 1.75 to 3.65 | 0.93 |
| | | | | | | | NRS average pain past week | 4.7 (2.53) | 2 (1.77) | 2.1 (1.82) 1.5 to 2.7 | 1.15 |
| No change - 2nd Peak Knee Adduction Moment | | | | | | | | | | | |
| Foroughi 2011a (Strengthening) | 0.07 | WOMAC function | 19.4 (9.8) | 13.3 (9.4) | 6.1 (7.08) 2.78 to 9.41 | 0.63 | WOMAC pain | 5.6 (3.2) | 3.8 (2.7) | 1.8 (2.46) 0.65 to 2.95 | 0.6 |
| Foroughi 2011a (Sham exercise) | 0.03 | WOMAC function | 23.3 (11.3) | 18.1 (12) | 5.2 (8.61) 1.65 to 8.75 | 0.44 | WOMAC pain | 5.6 (3.2) | 4.4 (3.7) | 1.2 (2.87) 0.02 to 2.38 | 0.34 |
| Hunt 2018 (Progressive walking) | 0.07 | WOMAC function | 21.4 (9.49) | 16.7 (9.47) | 7.7 (9.46) 4.5 to 10.9 | 0.81 | WOMAC pain | 6.4 (2.5) | 5.4 (3) | 1.5 (2.95) 0.5 to 2.5 | 0.51 |
| | | | | | | | NRS average pain past week | 3.7 (1.87) | 2.3 (1.8) | 1.8 (1.77) 1.2 to 2.4 | 1.02 |
| Changed – Peak Knee Abduction Moment | | | | | | | | | | | |
| Wang 2016 (Vibration + Quads strengthening) | -0.25 | WOMAC function | 35.7 (10.3) | 12.5 (9.7) | 23.2 (7.39) 19.64 to 26.76 | 2.31 | WOMAC pain | 13.2 (3.5) | 6.7 (2.8) | 6.5 (2.66) 5.22 to 7.78 | 2 |
| | | | | | | | VAS pain on joint movement | 8.2 (1.1) | 3.4 (0.9) | 4.8 (0.86) 4.38 to 5.22 | 4.69 |
| No change – Peak Knee Abduction Moment | | | | | | | | | | | |
| Wang 2016 (Quads strengthening) | -0.07 | WOMAC function | 33.8 (11.2) | 15 (8.2) | 18.8 (7.67) 15.21 to 22.39 | 1.8 | WOMAC pain | 12.7 (4) | 1.9 (0.5) | 10.8 (3.69) 9.07 to 12.53 | 2.4 |
| | | | | | | | VAS pain on joint movement | 8 (1.2) | 3.8 (0.9) | 4.2 (0.93) 3.77 to 4.63 | 3.82 |
| No change - Knee Flexion Range | | | | | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | 0.18 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Shen 2008 (Tai Chi) | 0.13 | WOMAC function | 40.6 (14.1) | 33 (13.2) | 7.6 (10.09) 4.37 to 10.83 | 0.55 | WOMAC pain | 16.3 94.3) | 13.2 (4.5) | 3.1 (3.62) 1.94 to 4.26 | 0.7 |
| | | | | | | | VAS overall pain previous week | 3.4 (2) | 3.1 (2.4) | 0.3 (2.06) -0.36 to 0.96 | No Change |
| | | | | | | | VAS Max pain | 5.2 (2.3) | 4.1 (2.8) | 1.1 (2.2) 0.39 to 1.81 | 0.42 |
| Changed - Peak Knee Flexion during Early Stance | | | | | | | | | | | |
| Zhu 2016 (Tai Chi) | 1.26 | WOMAC function | 24.7 (12.95 ) | 15.85 (7.6) | 8.85 (7.76) 5.32 to 12.39 | 1.14 | WOMAC pain | 8.46 (4.9) | 5.15 (3.24) | 3.31 (2.89) 2 to 4.63 | 1.15 |
| No change - Peak Knee Flexion during Early Stance | | | | | | | | | | | |
| DeVita 2018 (Strengthening) | 0 | WOMAC function | 17.7 (9.7) | 6.8 (9.5) | 10.9 (7.07) 6.32 to 15.14 | 1.13 | WOMAC pain | 6.43 (3.18) | 3.64 (2.55) | 2.79 (2.42) 1.45 to 4.13 | 0.95 |
| Fisher 1997 (Strengthening)^ | -0.17 | Jette Functional Index - Difficulty | 2.2 (0.8) | 1.9 (0.7) | 0.29 (0.59) -0.84 to 0.25 | No change | Jette Functional Index - Pain | 1.79 (0.53) | 1.40 (0.25) | 0.38 (0.41) 0.76 to 0.01 | 0.78 |
| Gaudreault 2011 (Exercise therapy) | 0.07 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Changed- Peak Knee Abduction Angle | | | | | | | | | | | |
| Wang 2016 (Vibration + Quads strengthening) | -0.41 | WOMAC function | 35.7 (10.3) | 12.5 (9.7) | 23.2 (7.39) 19.64 to 26.76 | 2.31 | WOMAC pain | 13.2 (3.5) | 6.7 (2.8) | 6.5 (2.66) 5.22 to 7.78 | 2 |
| | | | | | | | VAS pain on joint movement | 8.2 (1.1) | 3.4 (0.9) | 4.8 (0.86) 4.38 to 5.22 | 4.69 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Wang 2016 (Quads strengthening) | -0.53 | WOMAC function | 33.8 (11.2) | 15 (8.2) | 18.8 (7.67) 15.21 to 22.39 | 1.8 | WOMAC pain | 12.7 (4) | 1.9 (0.5) | 10.8 (3.69) 9.07 to 12.53 | 2.4 |
| | | | | | | | VAS pain on joint movement | 8 (1.2) | 3.8 (0.9) | 4.2 (0.93) 3.77 to 4.63 | 3.82 |
| No change - Peak Knee Abduction Angle | | | | | | | | | | | |
| Gaudreault 2011 (Exercise therapy) | -0.12 | WOMAC function | 24.9 (14.1) | 10.1 (9.7) | 14.84 (9.68) 11.16 to 18.52 | 1.13 | WOMAC pain | 7.92 (3.67) | 2.55 (2.66) | 5.37 (2.76) 6.52 to 9.32 | 1.6 |
| Changed - Maximum Sagittal Angular Velocity during Stance | | | | | | | | | | | |
| Roper 2013 (Aquatic treadmill) | 0.77 | NA | - | - | - | - | VAS pain past week | 37.2 (2.3) | 25.5 (25.2) | 15.4 (20.7) 3.45 to 27.35 | 0.74 |
| No change - Maximum Sagittal Angular Velocity during Stance | | | | | | | | | | | |
| Roper 2013 (Treadmill) | -0.4 | NA | - | - | - | - | VAS pain past week | 40 (24.1) | 37.4 (23.4) | 0.1 (19.2) -10.99 to 11.19 | No change |
| Changed - Minimum Knee Transverse Angular Velocity during Swing | | | | | | | | | | | |
| Roper 2013 (Aquatic treadmill) | 0.5 | NA | - | - | - | - | VAS pain past week | 37.2 (2.3) | 25.5 (25.2) | 15.4 (20.7) 3.45 to 27.35 | 0.74 |
| No change - Minimum Knee Transverse Angular Velocity during Swing | | | | | | | | | | | |
| Roper 2013 (Treadmill) | -0.54 | NA | - | - | - | - | VAS pain past week | 40 (24.1) | 37.4 (23.4) | 0.1 (19.2) -10.99 to 11.19 | No change |
| Changed - Acceleration Anterior-Posterior | | | | | | | | | | | |
| Turcot 2009 | 0.52 | NA | - | - | - | - | WOMAC pain | 196 (93.6) | 63.7 (66.4) | 137.77 (93) 93.03 to 171.57 | 1.42 |
| Changed - Knee Angle at Initial Contact | | | | | | | | | | | |
| Zhu 2016 (Tai Chi) | 0.89 | WOMAC function | 24.7 (12.95 ) | 15.85 (7.6) | 8.85 (7.76) 5.32 to 12.39 | 1.14 | WOMAC pain | 8.46 (4.9) | 5.15 (3.24) | 3.31 (2.89) 2 to 4.63 | 1.15 |

Change was defined as 95% confidence interval does not cross zero and standardised response mean >0.2.
Abbreviations: ADL, activities of daily living; CI, confidence interval; IQR, interquartile range; KOOS, Knee Injury and Osteoarthritis Outcome Scale; NA, not applicable; NRS, Numerical Rating Scale; PKAM, Peak knee adduction moment; SD, standard deviation; VAS, Visual Analogue Scale; WOMAC; Western Ontario McMaster Universities Osteoarthritis Index.

**Appendix 3-7. Direction of a change in movement parameter, activity limitation and pain (count (%)) for movement parameters that did not demonstrate a change.**

| | | Activity Limitation | | | Pain | | |
|---|---|---|---|---|---|---|---|
| | | Improved | No Change | Worse | Improved | No Change | Worse |
| | | | | Moments | | | |
| Knee flexion moment | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 10 (91%) | 1 (9%) | 0 | 11 (92%) | 1 (8%) | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee internal rotation moment | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 (100%) | 0 | 0 | 1 (100%) | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee external rotation moment | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| | | | | Kinematics | | | |
| Knee flexion range | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 4 (100%) | 0 | 0 | 1.66 (83%) | 0.33 (27%) | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee adduction angle | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 5 (100%) | 0 | 0 | 4 (80%) | 1 (20%) | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Max sagittal angular velocity | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 0 | 0 | 0 | 1 | 1 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee external rotation angle | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee angle swing | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 2 (100%) | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Peak knee internal rotation angle | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Acceleration medial-lateral | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 0 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Adduction-abduction range | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |
| Internal-external rotation range | Increased | 0 | 0 | 0 | 0 | 0 | 0 |
| | No Change | 1 | 0 | 0 | 1 | 0 | 0 |
| | Decreased | 0 | 0 | 0 | 0 | 0 | 0 |

Change was defined as 95% confidence interval does not cross zero and standardised response mean >0.2.

Abbreviations: EMG, surface electromyography; VL, vastus lateralis; VM, vastus medialis; BF, biceps femoris; ST, semitendinosus. * Vote counting system – 1.0 vote per parameter. Where multiple measures were used for an outcome, the 1 vote was divided by the total number times used. For example, if pain was reported using WOMAC pain and NRS pain, then the vote was divided equally, therefore if WOMAC pain changed and NRS pain did not change the score was 0.5. See Table 3-4 for movement parameters that did demonstrate a change.

## 3.10 Summary of Chapter 3

The aim of Chapter 3 was to systematically review cohort studies and randomised controlled trials to investigate how changes in knee joint movement parameters during functional activities relate to changes in activity limitation or pain after exercise intervention in people with knee osteoarthritis. A relationship between a change in a movement parameter and change in a clinical outcome (activity limitation or pain) occurred infrequently at a group level. Walking was the only activity that was biomechanically analysed despite negotiating stairs and transitioning to and from a chair also being recommended as part of the assessment of physical function in clinical guidelines. Limitations in methodology and participant characteristics within the included studies may have influenced the results and will be explored within Chapter 7 (the Discussion). Despite the findings of the systematic review, there is some evidence in low back pain research that suggests that appropriate methodological design and individualised assessment of participant characteristics may result in stronger relationships between change in movement patterns and clinical outcomes after targeted intervention. Considering there are similarities in biopsychosocial factors between people with low back pain and knee osteoarthritis, a similar finding might be potentially present in the population with knee osteoarthritis.

# Chapter 4

# Study 2a: Human Activity Recognition

## Human Activity Recognition for People with Knee Osteoarthritis: A Proof-of-Concept

Assessing physical function outcomes is a core component for the guideline-based management of a person with knee osteoarthritis. Patient-reported outcome measures were selected as a measure of physical function in the systematic review presented in Chapter 3. Physical function is a multidimensional construct that includes aspects related to patient perception of function (patient-reported outcome measures) as well as objective performance of activities (performance tests and activity monitoring).

Activity monitoring is one method of collecting data about actual physical performance of activities in free-living environments (such as work, home and in the community). The most common form of activity monitoring is through the use of accelerometers to measure energy expenditure. But this method of activity monitoring does not provide information about the particular activities that are being performed.

IMU-based human activity recognition provides an opportunity to quantify performance of clinically relevant activities in free-living environments. Yet, at the time of this research, no studies had validated IMU-based human activity recognition systems on data collected from people who have knee osteoarthritis, an essential step prior to using such a system in free-living environments.

This chapter presents a proof-of-concept study that explored the development of a human activity recognition system for people with knee osteoarthritis. The aim of this study was to explore the feasibility of using IMU data collected from people with knee osteoarthritis to train a machine learning model to classify clinically important

activities and phases of those activities. The activities selected align with guideline recommendations for the assessment of physical function in people with knee osteoarthritis, and the phases were selected as part of an algorithmic data handling pipeline that could be potentially used for subsequent biomechanical analysis.

**This chapter was published in the journal *Sensors*.**

**Tan, J.-S.**, Beheshti, B. K., Binnie, T., Davey, P., Caneiro, J. P., Kent, P., Smith, A., O'Sullivan, P., & Campbell, A. (2021). Human Activity Recognition for People with Knee Osteoarthritis—A Proof-of-Concept. *Sensors*, *21*(10). https://doi.org/10.3390/s21103381

## 4.1 Abstract

Clinicians lack objective means for monitoring if their knee osteoarthritis patients are improving outside of the clinic (e.g. at home). Previous human activity recognition models using wearable sensor data have only used data from healthy people and such models are typically imprecise for people who have medical conditions affecting movement. Human activity recognition models designed for people with knee osteoarthritis have classified rehabilitation exercises but not the clinically relevant activities of transitioning from a chair, negotiating stairs and walking, which are commonly monitored for improvement during therapy for this condition. Therefore, it is unknown if a human activity recognition model trained on data from people who have knee osteoarthritis can be accurate in classifying these three clinically relevant activities. Therefore, we collected inertial measurement unit (IMU) data from 18 participants with knee osteoarthritis and trained convolutional neural network models to identify chair, stairs and walking activities, and phases. The model accuracy was 85% at the first level of classification (activity), 89% to 97% at the second (direction of movement) and 60% to 67% at the third level (phase). This study is the first proof-of-concept that an accurate human activity recognition system can be developed using IMU data from people with knee osteoarthritis to classify activities and phases of activities.

## 4.2 Introduction

Osteoarthritis is one of the leading causes of disability (Ackerman et al., 2019; Vos et al., 2012). People with knee osteoarthritis have symptoms such as pain and stiffness that result in difficulty performing specific physical activities such as transitioning from a chair, negotiating stairs (Machado et al., 2008) and walking (Wilkie et al., 2007). A recent review on the application of machine learning for people with knee osteoarthritis identified that movement-based (biomechanical) data has predominantly been used for the purposes of diagnosis and prediction of outcome in people with knee osteoarthritis (Kokkotis et al., 2020). While there is growing interest in wearable sensor technology for use in clinical environments, no studies have investigated if machine learning approaches can assist with monitoring

improvement in the performance of clinically relevant activities outside of a clinical environment, for patients diagnosed with knee osteoarthritis (Kokkotis et al., 2020).

There is high-quality evidence of improvements in pain and function following movement interventions, such as exercise, or surgical interventions. For people receiving these treatments, the leading medical society dedicated to researching osteoarthritis, the Osteoarthritis Research Society International, recommends that people who have a confirmed diagnosis of knee osteoarthritis are monitored for improvement in the performance of three specific and clinically relevant everyday activities; (a) transitioning from a chair, (b) negotiating stairs and (c) walking (Dobson et al., 2013). These movements are clinically relevant because they are related to pain, stiffness, and reduced ability to participate in society.

To assess if someone diagnosed with knee osteoarthritis is improving because of a treatment, clinicians are currently limited to assessing the performance of clinically relevant painful activities only in observed conditions such as in a clinic. Usually, a clinician would watch their patient perform the activity, however a single observation in a clinic does not demonstrate if a person is avoiding or doing less of this activity when unobserved after leaving the clinic.

Currently, the best method of assessing unobserved activities is to use questionnaires known as patient-reported outcome measures (Dobson et al., 2013). However, questionnaires can be unreliable because they assess a patient's perception of their ability to perform an activity which does not objectively measure how many times or how they actually perform an activity when at home or at work. Wearable sensor technology has the potential to be used to help clinicians monitor the patient's progress when they are unobserved.

One type of wearable sensor, inertial measurement units (IMUs), can collect movement-based information that can be processed into clinically relevant biomechanical data through fusion (Weygers et al., 2020) or machine learning algorithms (Mundt, Koeppe, David, Witter, et al., 2020). These methods have been reported to provide biomechanical outputs that are useful for clinicians, such as kinematics (e.g. knee flexion angle or knee angular velocity) (Drapeaux & Carlson,

2020; Mundt, Koeppe, David, Witter, et al., 2020; van der Straaten et al., 2018) during a specific phase of an activity (e.g. stance phase of ascending stairs) when the patient is observed as part of an assessment in the clinic.

IMUs enable data collection outside of the clinical environment and can be used when patients are unobserved. However, IMUs create large and unlabelled datasets making it difficult to identify which activities were performed, because the patient was unobserved. One approach to identifying when an activity or phase of an activity was performed from IMU data is through a machine learning approach known as human activity recognition. Human activity recognition models are built from algorithms to automate the process of classifying performance of human activities. As there is limited feasibility and practicality for clinicians to observe people with knee osteoarthritis in unobserved conditions, human activity recognition has the potential to provide clinically relevant activity data from large continuous datasets. Having a system that can automatically label when an activity was performed could be subsequently used to monitor if a patient is improving by providing objective data about whether they are performing an activity more frequently, or to segment the data so that it can be used for subsequent biomechanical analysis.

The majority of human activity recognition models are built using traditional machine learning approaches (e.g. support vector machines, random forest, k-nearest neighbour); however, alternative approaches using deep neural networks such as convolutional neural networks (CNN) have recently demonstrated superior accuracy (Brock et al., 2017; Chen & Xue, 2015; Fridriksdottir & Bonomi, 2020; Jiang & Yin, 2015). The benefit of a CNN model is that it automatically detects important features from input data, minimising programming requirements typically required for traditional machine learning approaches. In addition, deep learning approaches like CNN are able to handle nonlinear interactions between features, something which is limited when using traditional machine learning approaches where features are defined by the researcher. There are many laboratory studies that have reported the accuracy of classifying physical activities from IMU data (Arif & Kattan, 2015; Ascioglu & Senol, 2020; Charlton et al., 2017; Chen et al., 2015; Chen & Xue, 2015; Cust et al., 2019; Emmerzaal et al., 2020; Fridriksdottir & Bonomi, 2020; Hendry et

al., 2020; Huang et al., 2017; Jiang & Yin, 2015; Martinez-Hernandez & Dehghani-Sanij, 2018, 2019; O'Reilly et al., 2018; Rast & Labruyère, 2020; Whiteside et al., 2017) using both traditional and deep learning approaches. Ramanujam et al. (2021) provide a review of the most up-to-date computational advances in deep learning for human activity recognition which is beyond the scope of this paper.

There are multiple studies that have developed human activity recognition models to classify daily activities using the lower limbs (e.g. standing, walking, going up stairs, walking down a hill), with accuracy ranging from 83% to 98% using training data from healthy people (Arif & Kattan, 2015; Ascioglu & Senol, 2020; Charlton et al., 2017; Chen et al., 2015; Cust et al., 2019; Emmerzaal et al., 2020; Fridriksdottir & Bonomi, 2020; Hendry et al., 2020; Jiang & Yin, 2015; O'Reilly et al., 2018). Other studies have classified specific phases of activities like transitioning from sit to stand or the stance phase of walking, reporting accuracy >99% (Martinez-Hernandez & Dehghani-Sanij, 2018, 2019). In studies that have specifically developed human activity recognition models for people who have knee osteoarthritis, models have been reported to classify a variety of rehabilitation exercises with accuracy >97% (Chen et al., 2015; Huang et al., 2017). These studies are limited for two reasons.

The first limitation is that these models have been trained and tested on healthy, typically young, participants rather than people with knee osteoarthritis (Arif & Kattan, 2015; Ascioglu & Senol, 2020; Chen et al., 2015; Emmerzaal et al., 2020; Fridriksdottir & Bonomi, 2020; Huang et al., 2017; Martinez-Hernandez & Dehghani-Sanij, 2018, 2019; O'Reilly et al., 2018). These human activity recognition models have not yet been trained and tested on people who have been diagnosed with knee osteoarthritis. There is a substantial body of research demonstrating that movement characteristics of people who have knee osteoarthritis are significantly different from those of healthy people when performing activities like transitioning from a chair, negotiating stairs and walking (Astephen et al., 2008; Baliunas et al., 2002; Iijima et al., 2018; Turcot et al., 2012). There is a high level of movement pattern variability across the population with knee osteoarthritis which is affected by structural severity (Astephen et al., 2008), gender (Kiss, 2011), perceived instability (Gustafson et al., 2015). Therefore, it is currently unknown if a human activity recognition model could

accurately classify activities from training data collected from people who have knee osteoarthritis.

Previous studies have demonstrated that human activity recognition models trained on data from healthy people are less accurate for use in people who have health conditions that affect how they move (Albert et al., 2012; Lonini et al., 2016). For example, using a support vector machine model trained on data from healthy people, Albert et al. (2012) reported significantly less accuracy of their model when tested on people who have Parkinson's disease (75%), compared to healthy participants (86%). In another study using a random forest classifier, Lonini et al. (2016) reported a median accuracy that was 26% lower for classification predicting five activities when training data using healthy participants was tested on people who use knee–ankle–foot orthoses due to lower limb impairments. Together, these studies suggest that human activity recognition models trained on data from people who have abnormal movement characteristics because of medical conditions (e.g. knee osteoarthritis or Parkinson's disease) result in poorer test accuracy in the patient population.

The second limitation is that so far, studies that have developed human activity recognition for people who have knee osteoarthritis have only trained and tested human activity recognition models to classify rehabilitation exercises in healthy people (Chen et al., 2015; Huang et al., 2017; O'Reilly et al., 2018) rather than functional activities such as walking, standing from a chair or using stairs, which are activities most important for clinicians to monitor for improvement after being prescribed exercises or after surgery.

To date, no studies have addressed these two limitations when reporting the development and validation of a human activity recognition model intended for use in people with knee osteoarthritis. Therefore, we aimed to develop a human activity recognition system that could classify the activities and phases of transitioning from a chair, negotiating stairs and walking using raw IMU training data from people who have knee osteoarthritis using CNN models. In this study, we have demonstrated a proof-of-concept that data collected from people who have a confirmed diagnosis of knee osteoarthritis can feasibly be used to train a human activity recognition model

using CNN architecture to classify clinically relevant activities and phases of activities at an acceptable level of accuracy.

## 4.3 Materials and Methods

Eighteen participants with the clinical diagnosis of knee osteoarthritis (National Institute for Health & Care Excellence, 2014) were recruited from local health provider clinics through direct referral and noticeboards. Inclusion criteria included ≥3 months of pain, ≥4/10 pain on most days and moderate activity limitation (a single item on the Function, Daily Living subscale of the Knee injury and Osteoarthritis Outcome Score) (Roos & Lohmander, 2003). Exclusion criteria were previous lower limb arthroplasty, severe mobility impairments (e.g. neurological disorders, fracture) or an inability to complete the physical assessment due to language or cognitive difficulties. As soft tissue artefacts result in 'noise' in IMU data, to minimise the impact of this, participants were excluded if they had a body mass index (BMI) > 35 kg/m2 or >30 kg/m2 and relatively more soft tissue around the thigh with a waist-to-hip ratio (WHR) of ≤0.85 for women and ≤0.95 for men. All participants provided written informed consent and institutional ethics approval was obtained (HRE2017-0695) prior to data collection. The characteristics of participants are reported in **Table 4-1**.

**Table 4-1. Characteristics of participants**

| Characteristics | Mean (SD) | Range |
|---|---|---|
| Age (years) | 66.2 (8.7) | 49 to 82 |
| Female (%) | 53% | |
| Weight (kilograms) | 80.5 (15.9) | 44 to 113 |
| Height (metres) | 1.7 (0.1) | 1.6 to 1.9 |
| Body mass index (kg/m$^2$ | 26.6 (15.9) | 17.8 to 33.4 |

### 4.3.1 Data Collection

Data were collected during a single session (average approximately 30 min) in a motion analysis laboratory. Height and weight data were collected using a manual stadiometer and calibrated digital scale prior to placing IMUs (v6 research sensors, DorsaVi, Melbourne, Australia) and retro-reflective markers on the participant. Participants then performed flexion–extension of the knee approximately 10 times as

warm-up movements for both knees. A standardised battery of functional activities was performed by participants that included: transitioning from a chair (5 trials of sit-to-stand-to-sit on wooden box 40 cm in height), negotiating stairs (3 trials of a 3-stair ascent, 3 trials of a 3-stair descent with each step 20 cm in height) and walking (3 trials of a 5 m self-paced walk).

### 4.3.2 Activities for Classification

The levels of classification are outlined in **Table 4-2**. Three activities were classified at the first level, four at the second level and six at the third level.

**Table 4-2. Levels of Classification**

| Level 1 | Level 2 | Level 3 |
|---|---|---|
| Chair | Sit down<br>Stand up | |
| Stairs | Stairs ascending | Stance<br>Swing |
| | Stairs descending | Stance<br>Swing |
| Walking | | Stance<br>Swing |

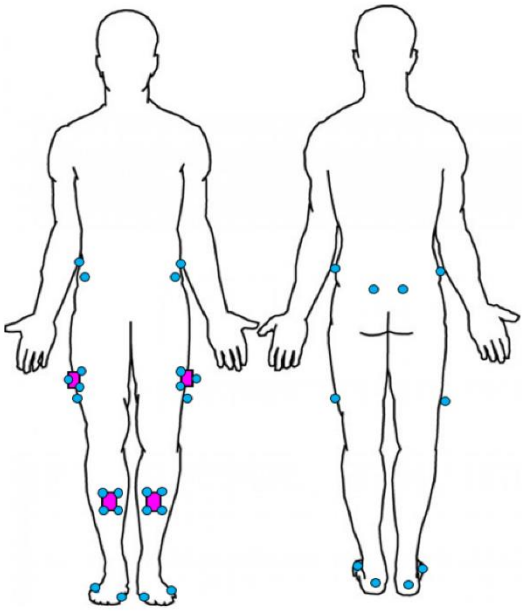
### 4.3.3 Instrumentation

Two data collection systems were used. To train the human activity recognition model, we used data from four DorsaVi IMUs placed on the thighs and shanks (**Figure 4-1**) of the participants. The IMUs included a triaxial accelerometer and gyroscope, weighed 17 grams and measured 4.8 x 2.9 x 1 centimetres. The location and number of sensors selected are the minimum number required to enable subsequent biomechanical analysis of both knees (e.g. knee joint flexion angle).

To precisely label the start and end times of each trial, we used a second system, an 18 camera Vicon three-dimensional motion analysis system (Oxford Metrics Inc., Oxford, UK). The events defining the start and end times are described in **Appendix 4-1**. This was required as the IMU data were not able to be directly labelled while collecting data in the laboratory. Therefore, we time-synchronised the IMU and Vicon systems to allow labelling of the start and end of each trial in the IMU data.

Synchronisation procedures were performed by placing IMUs in a wooden box (with three retro-reflective markers attached) and rotated 10 times, >90° around the IMU's X axis during a single Vicon data collection trial.

**Figure 4-1. Placement of IMUs (purple) used for training the CNN models and Vicon marker (blue) placement for recording start and end times for each trial.**



Bilateral Retroreflective marker locations:

- Anterior superior iliac spine
- Posterior superior iliac spine
- Superior iliac crest
- Lateral thigh between greater trochanter and lateral epicondyle 1-2cm below IMU
- 1st and 5th metatarsal head
- Calcaneal tuberosity
- 3 point cluster over thigh IMUs
- 4 point cluster over tibia IMUs

Bilateral IMU locations:

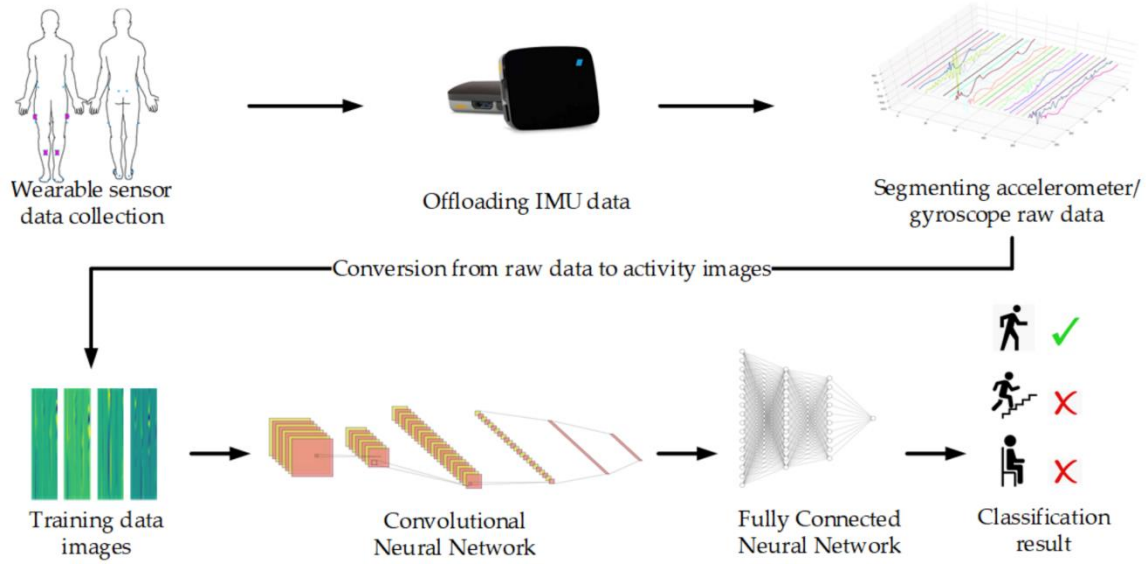
- Lateral thigh 2/3 distal to greater trochanter
- Anteromedial tibia 1/2 between tibial tuberosity and talocrural joint line

The sampling frequency for the IMU and Vicon systems were 100 and 250 Hz, respectively. IMUs were placed in a standardised manner according to manufacturer instructions by an experienced musculoskeletal physiotherapist bilaterally on the lower limbs with double-sided hypoallergenic tape halfway between the superior edge of the greater trochanters and lateral epicondyles, halfway between the tibial tuberosities and anterior talocrural joints **Figure 4-1**. Twenty-eight retro-reflective markers for Vicon motion analysis were placed on anatomical landmarks of the pelvis and lower limb (**Figure 4-1**) consistent with previously published models which align with the International Society of Biomechanics recommendations (Wu et al., 2002).

### 4.3.4 Human Activity Recognition System Development

The architecture of the human activity recognition system we developed is detailed through sections 4.3.4.1 to 4.3.4.3 and summarised in **Figure 4-2**. Further details about the CNN and fully connected network are in **Appendix 4-2**.

**Figure 4-2. Architecture of the proposed human activity recognition system.**




#### 4.3.4.1 Data Preparation

Raw IMU data (tri-axial accelerometer, gyroscope, and magnetometer) were offloaded and output as time stamped files for each sensor. The reference standard kinematic data were processed in Vicon Nexus software (Oxford Metrics Inc., Oxford, UK). Reconstructed Vicon quaternion data and the filtered raw orientation data from each sensor's accelerometer, gyroscope and magnetometer were time-synchronised by use of normalised cross-correlation using a customised LabVIEW program (National Instruments, Austin, TX, USA). Start and end times were exported for each activity for the raw IMU and reconstructed Vicon data. As there were a different number of samples for the stair phases (swing/stance), this resulted in an unbalanced dataset, which reduced model accuracy because of overfitting. Therefore, we balanced the dataset with an automated randomisation procedure, whereby the dataset was shuffled each time and the number of selected samples was balanced to optimise the accuracy of the model. Magnetometer data were then discarded as it was not required for the development of the human activity recognition model.

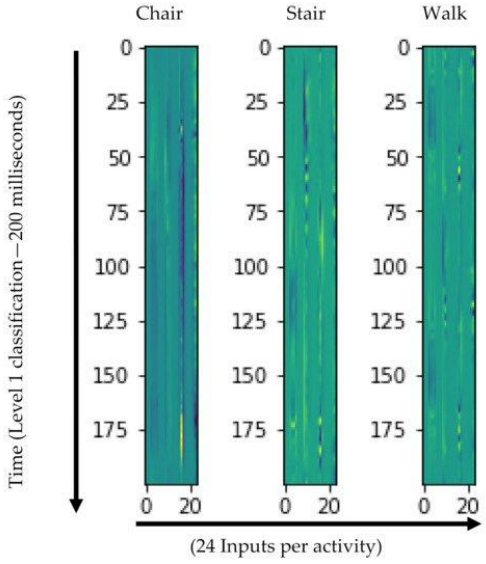
#### 4.3.4.2 Classification

One contemporary method for human activity recognition model development is a machine learning approach known as deep learning. Deep learning uses a

programmable neural network that automatically learns classification features from raw data, reducing the programming requirements used for other traditional machine learning methods. This approach automatically identifies complex features, rather than using predefined time and frequency domain features required for traditional machine learning human activity recognition model approaches, such as support vector machines or k-nearest neighbour (Hou, 2020; Kautz et al., 2017; Sani et al., 2017; Wu et al., 2002). A convolutional neural network (CNN) is one type of deep learning approach that can be used for high dimensional time-series data (LeCun et al., 2015) that outperforms traditional machine learning approaches (Brock et al., 2017; Chen & Xue, 2015; Fridriksdottir & Bonomi, 2020; Jiang & Yin, 2015).

Deep neural networks such as CNNs are ideal for handling image data, like those from IMUs time-series data, which can be arranged into a two-dimensional 'image' as an input matrix. Features are then extracted automatically as each activity ideally represents a unique activity 'image' pattern (**Figure 4-3**). Input included triaxial (x, y, z) accelerometer and gyroscope data from all four IMUs resulting in a total of 24 inputs. Each input was stacked column by column then segmented into fixed size windows according to the level of classification. The images were then converted from a segmented numerical data array of the 12 accelerometer and 12 gyroscope inputs into images by normalising the dataset to 0 to 255 range required for digital image production.

**Figure 4-3. Visual representation of Level 1 activity 'image' patterns.**

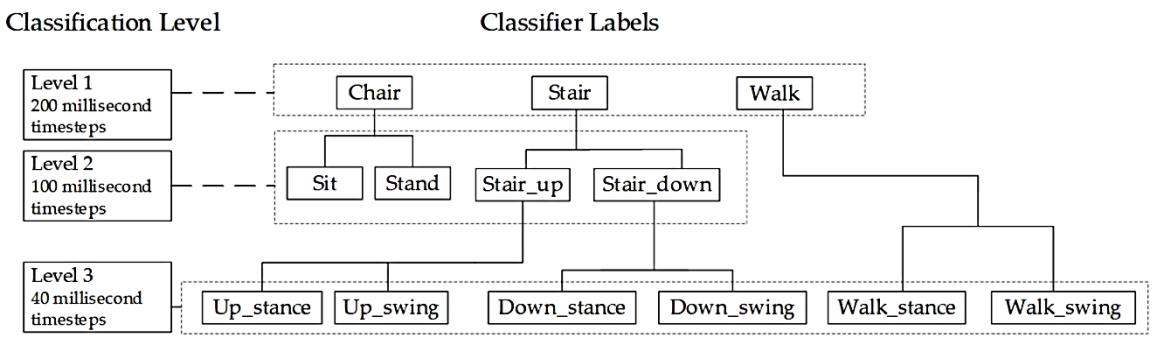


Error! Reference source not found. depicts the CNN model architecture for each level of classification. Various architectures were tested using a different number of CNN layers, number of neurons in each layer, activation function, CNN kernel size, number of filters, max pooling size, number of dense layers in the fully connected part of the model, learning rate for 'Adam' optimisation, and number of epochs. To avoid overfitting, we used cross-validation (see section 4.4), data augmentation, adjusted the learning rate and used dropout and early stopping functions. Automatic feature extraction was performed at the first level of classification using two convolutional layers and three for the second and third level of classification. We developed a total of six models: one for level 1 (to classify between chair, stairs, and walking), one for level 2-Chair (to classify between Sit down and Stand up), one for level 2-Stairs (to classify between Stairs ascending and Stairs descending), one for level 3-Stairs-Stairs ascending (to classify between Stance and Swing), one for level 3-Stairs-Stairs descending (to classify between Stance and Swing), and finally one for level 3-Walking (to classify between Stance and Swing). An optimisation algorithm (Adam) trained the model (Kingma & Ba, 2014) using different learning rates depending on the activity.

#### 4.3.4.3 Segmentation

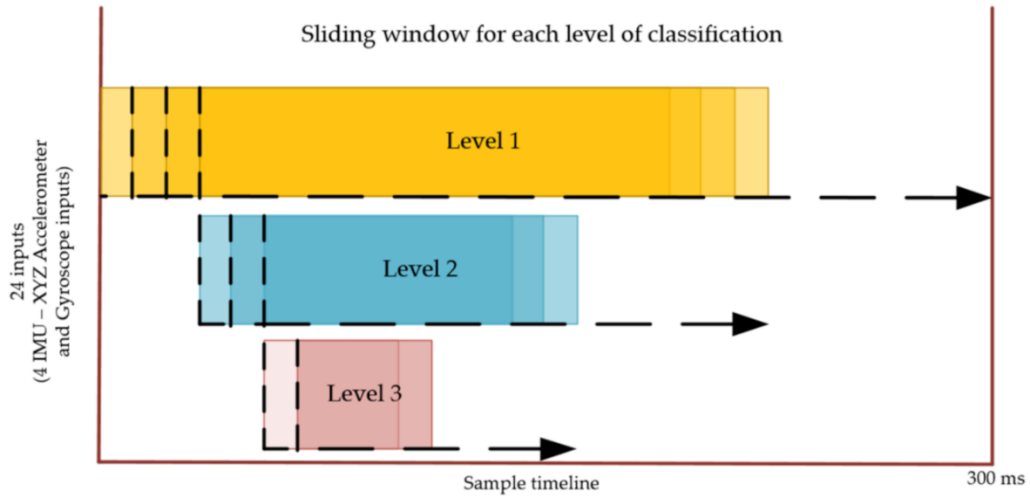
A decision tree was developed (**Figure 4-4**), with a separate model created for each activity. IMU data were segmented at fixed window sizes of 200, 100 and 40 milliseconds for each subsequent level of classification (**Figure 4-5**). Each of these windows slid 10 milliseconds over the trial.

**Figure 4-4. Decision tree for three levels of activity classification – Level 1 Activity, Level 2 Direction, Level 3 Phase.**



The fixed window size was optimised to train the model by testing multiple window sizes for each level of classification. To determine the best window size, we created a distribution graph for all the trials and selected the window size based on the 80$^{th}$ percentile. An additional training image was produced for each additional 10 milliseconds where the duration was longer than the fixed window size. For example, if the trial length was 240 milliseconds, then there were five images created which allowed data augmentation. In situations where there was an overlapping window between two activities (e.g. Level 3 classification between swing and stance), the window was classified based on the higher prediction probability.

**Figure 4-5. Illustration of the segmented window sliding in 10 ms increments for each level of classification. Level 1—200 ms; Level 2—100 ms; Level 3—40 ms.**




## 4.4 Model Performance—Statistical Testing

The accuracy of the models was evaluated using a leave-one-subject-out cross-validation (LOSOCV) method (Gholamiangonabadi et al., 2020). This method of validation trains the model on all participants except one and independently tests the model on the participant that is 'left out' (e.g. trained on 17, tested on one). Tests are repeated until each participant has been left out, and the reported result is the average across all participants. This model was chosen as it is more clinically relevant, as LOSOCV provides an estimate that would more closely approximate the average accuracy for individual patients than other validation approaches (Gholamiangonabadi et al., 2020). We evaluated the accuracy (1), precision (2) and recall (3) of our model. In addition, we present a confusion matrix which depicts the

total of the binary (correct/incorrect) classifications from LOSOCV across all participants.

1. $$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative}$$

2. $$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

3. $$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

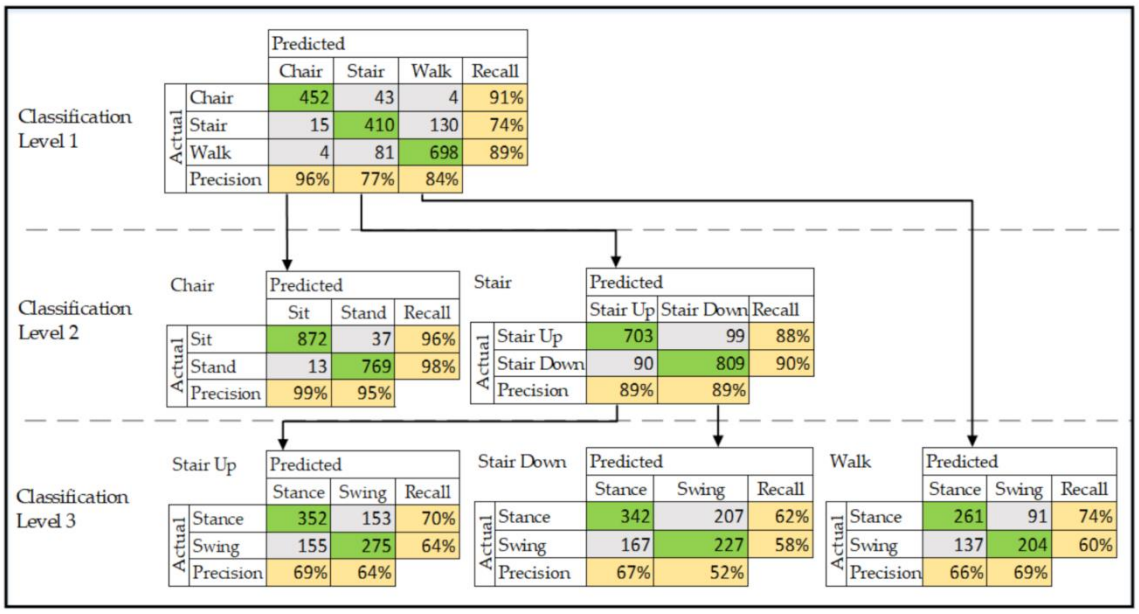
## 4.5 Results

All participants completed each activity per protocol. The overall accuracy across multiple levels of classification ranged from 60% to 97% (**Table 4-3**). Confusion matrices are presented in **Figure 4-6** for each level of classification.

**Table 4-3. Accuracy of CNN models using leave-one-out cross-validation for each level of classification.**

| | Accuracy |
|---|---|
| 1st Level Classification | |
| Chair, Stair, Walk | 85% |
| 2nd Level Classification | |
| Chair - stand/sit | 97% |
| Stair - up/down | 89% |
| 3rd Level Classification | |
| Stair up stance/swing | 67% |
| Stair down stance/swing | 60% |
| Walk stance/swing | 67% |

**Figure 4-6. Confusion matrices for classification of activities/phases per classification level.**




Green cells represent correct classification and arrows represent the classification pathway from activities to phases of activities. Yellow cells represent the prediction accuracy.

## 4.6 Discussion

Previous literature reporting the development of human activity recognition models for people with knee osteoarthritis had not explored (a) the potential accuracy of a model trained on data collected from people who have knee osteoarthritis rather than healthy people, and (b) the capacity of such a model to classify activities related to the disability experienced by people with knee osteoarthritis rather than rehabilitation exercises. Therefore, the aim of this proof-of-concept study was to use IMU data collected from people who have knee osteoarthritis to train a human activity recognition system to classify clinically relevant activities and phases of those activities.

While many previous human activity recognition studies have investigated novel computational methods to optimise human activity recognition models, we took a different approach and demonstrated two novel findings, that (a) a human activity recognition model can be trained on IMU data collected from participants who have knee osteoarthritis rather than healthy people, and (b) activities of transitioning from a chair, negotiating stairs and walking can be classified from training data collected

from this specific population. The model accuracy was 85% at the first level of classification, 89% to 97% at the second and 60% to 67% at the third. Our model performed with a high degree of accuracy compared to studies using the same validation approach (LOSOCV).

### 4.6.1 Comparison of Human Activity Recognition System Accuracy to Previous Literature

The results of our human activity recognition model (accuracy range 60% to 97%) are consistent with previous studies that classified activities of the lower limb in healthy people and people with medical conditions that affect their movement (e.g. Parkinson's disease) which have reported model accuracy between 75% and 99% (Albert et al., 2012; Arif & Kattan, 2015; Ascioglu & Senol, 2020; Deep & Zheng, 2019; Emmerzaal et al., 2020; Fridriksdottir & Bonomi, 2020; Gholamiangonabadi et al., 2020; Lonini et al., 2016; Nguyen et al., 2017). These previous studies use a single human activity recognition model to classify between 5 and 12 activities resulting in an overall accuracy for the single model. As our human activity recognition system used a decision tree framework to classify both activities and phases of those activities for clinical purposes that are described in section 4.6.4, comparison of our model is limited because we developed multiple human activity recognition models for three levels of classification. Those three levels resulted in a total of six models with accuracy reported for each separate model (**Figure 4-6**). However, and despite this, the accuracy for the first and second level of classification (85% to 97%) are promising compared to previous studies that used a single model.

At subsequent levels of classification, our model's accuracy slightly improved at the second level of classification (range 89% to 97%), but at the third level of classification, swing and stance phases for stairs and walking, the accuracy (range 60% to 67%) was reduced. Usually, the accuracy of human activity recognition algorithms reduces for finer levels of classification as predictions become more complex. Our results are consistent with these previous studies, where reductions in accuracy for subsequent level classification have been reported to range between 4.2% and 16.4% (Hendry et al., 2020; Whiteside et al., 2017). While we compared the results of our models to previous human activity recognition literature, the considerable

heterogeneity between these studies reduces the capacity to make meaningful comparisons (O'Reilly et al., 2018). This heterogeneity arises from a wide variety of factors that include: the type of activities to be classified, the number and location of sensors, the number of activities, the number of training data samples (and participants), the population sampled (e.g. healthy, knee osteoarthritis, Parkinson's disease) and the validation approach (e.g. LOSOCV or k-fold cross-validation).

The model was more accurate when classifying activities where movement patterns are distinct. For example, at the first level of classification, a chair transition was most frequently classified correctly (recall 91% - 452 correct predictions from 499 observations). We believe the higher accuracy for a chair transition is due to a unique activity pattern where both legs perform synchronised movements. On the other hand, negotiating stairs was misclassified as walking 23% of the time (130 incorrect predictions from 555 observations). As walking and negotiating stairs share similar features of a reciprocal movement pattern where one leg is swinging forward while the other is in stance moving backwards alternating in a rhythmical manner, misclassification between these two activities is commonly reported in studies that trained human activity recognition models using data from healthy participants (Ascioglu & Senol, 2020; Chen & Xue, 2015; Fridriksdottir & Bonomi, 2020). One solution recently reported is to combine neural network models to classify eight activities that included walking, walking uphill, walking down hill, ascending stairs, descending stairs and running – all activities that share a reciprocal pattern (Ascioglu & Senol, 2020). Combining CNN with another deep learning approach known as long short-term memory resulted in superior model performance with fewer misclassifications for activities that have reciprocal patterns compared to an independent CNN or long short-term memory models alone.

### 4.6.2 Appropriate Validation Approaches for Clinical Populations

The most important consideration when comparing the accuracy of a human activity recognition model is the validation approach. For IMU human activity recognition systems to be widely adopted and accepted by healthcare clinicians and researchers, the validation approach described in the machine learning field is very important. It is important for a clinician to know the average error that exists for

individual patients. For any particular patient where new data will be tested against the model, clinicians can have greater confidence in the accuracy of models that are validated with LOSOCV rather than other validation approaches (such as k-fold or 70:30 cross-validation) (Gholamiangonabadi et al., 2020).

Methods other than LOSOCV inflate the accuracy of models designed to be used on a single individual. For example, one study (Janidarmian et al., 2017) aggregated multiple human activity recognition datasets that represented multiple machine learning models, sensor types and placements that included ambulatory activities such as walking, ascending and descending stairs, and jogging, amongst others. They reported an accuracy of 96.4% when using a 10-fold cross-validation compared with a 79.9% using LOSOCV, representing a substantial 16.5% difference between these validation methods (Janidarmian et al., 2017). Similarly, in a human activity recognition model that included walking and multiple stationary activities in a sample of participants with Parkinson's disease, the accuracy of a support vector machine classifier reduced from 92.2% using a 10-fold cross-validation to 75.1% when using a LOSOCV (Albert et al., 2012). Therefore, in comparison, the accuracy of our CNN models ranging from 85% to 97% at the first and second level of classification using LOSOCV is promising.

Although LOSOCV has lower reported accuracy than some other validation approaches, it is preferred as it accounts for between-participant variability (Albert et al., 2012; Gholamiangonabadi et al., 2020; Janidarmian et al., 2017) which is important for a clinician who needs to know the average accuracy of a model as it applies to each new individual patient.

### 4.6.3 The Importance of Representative Sampling

Validating a model for populations who have movement impairments related to medical conditions is particularly important because they move differently from healthy people. In the introduction of this paper, we provided a detailed description of the importance of training human activity recognition models on data collected from people with medical conditions that affect their movement rather than healthy people. Briefly, previous studies have demonstrated that human activity recognition models trained on data from healthy people are less accurate when tested on data

collected from people with medical conditions that affect the way they move, reducing the model accuracy by up to 28% (Albert et al., 2012; Lonini et al., 2016).

### 4.6.4 Clinical Application of Human Activity Recognition

A central component of an initial clinical interaction is to establish a diagnosis. After a person is diagnosed with knee osteoarthritis, they should be referred for core interventions like movement rehabilitation or surgery because both these treatments are helpful to improve pain and ability to perform activities like transitioning from a chair, negotiating stairs and walking. However, currently clinicians do not have an objective method to use outside of the clinical environment to monitor if people are improving after treatment. Therefore, a human activity recognition system that can classify both clinically relevant functional activities and phases of those activities is potentially important for a clinician because it could help provide information about whether a person who has knee osteoarthritis is improving when outside of the clinic such as when at home or at work.

There are two potential ways a human activity recognition system could be used for clinical purposes for a person with knee osteoarthritis when they are unobserved (at home or at work) while under the care of a clinician. Firstly, for the purposes of physical activity monitoring. For instance, one patient with knee osteoarthritis may avoid using the stairs due to a fear of falling. The clinician's goal may be to increase use of stairs and therefore they could use a human activity recognition system to automatically and objectively count the number of times their patient used stairs during a period of physical activity monitoring. This is especially important as patient self-report of physical activity does not consistently correlate with wearable sensor-based monitoring (Jasper et al., 2021; Kowalski et al., 2012).

Secondly, a human activity recognition-IMU system could be used to segment the data for subsequent biomechanical analysis of activity phases for data collected when unobserved, outside of the clinic. Currently, most IMU systems can only be used in observed conditions where the start and end times of a data capture trial are known. Under unobserved conditions, for instance when a person is at home or at work, IMU data are unlabelled which currently requires a clinician to process long, continuous datasets which is time-consuming and therefore not feasible. Human

activity recognition provides a solution to segment the data for unobserved biomechanical analysis. For instance, a patient may have specific difficulties with the stance phase of descending stairs because of their stiff knee. The clinician's goal in this situation may be to change specific biomechanical patterns with the aim of reducing knee stiffness during this activity. In this case, the clinician may use human activity recognition to identify a window in the IMU data when the stance phase occurred when walking down stairs for subsequent biomechanical analysis, to monitor if their patient can bend their knee more when going down stairs after they receive treatment. Therefore, a human activity recognition system that can automatically segment the data by labelling the start and end times of a phase of an activity, when the patient is unobserved, could help a clinician monitor improvement of a specific movement parameter (e.g. knee joint angle or force).

### 4.6.5 Clinician and Patient Burden

It is important that human activity recognition model development considers both the intended population that will wear the sensors (e.g. a patient with knee osteoarthritis) as well as the intended population that will use that information (e.g. a clinician). Therefore, the choice as to the number of IMUs should balance the patient burden of wearing multiple sensors with optimising the accuracy of the model and any potential clinical benefit that may ensue. We chose to use a total of four IMUs on each participant, with two sensors on each lower limb for two reasons.

Firstly, studies have demonstrated higher accuracy with more sensors, especially across multiple body regions (Hendry et al., 2020; Lee et al., 2020). For example, using a 10-fold cross-validation, Lee et al. (2020) demonstrated a reduction in human activity recognition model accuracy of up to 8% when using inputs from three, rather than five IMUs when classifying six different squatting tasks (Janidarmian et al., 2017). Secondly, a sensor placed on both the thigh and the shank is required for subsequent biomechanical analysis of the knee using current IMU fusion algorithms. We therefore believe that four IMUs strikes a balance between optimal accuracy and participant burden while providing clinically relevant information for knowing if a patient is improving during treatment.

Studies have determined the best location for IMU placement that optimises model accuracy from as few as one sensor in healthy people (Janidarmian et al., 2017). However, optimal sensor positions could differ in healthy people compared to those with medical conditions that affect their movement. Additionally, the purpose of the human activity recognition model needs to be considered, especially if biomechanical analysis is required which requires at least two sensors to estimate the movement parameters around a joint like the knee. Future human activity recognition model development for people who have medical conditions that affect their movement, such as people who have knee osteoarthritis, should carefully consider IMU placement for the two clinical purposes of (a) identifying activities to count the frequency or duration of performance, or alternatively, (b) as a means of segmenting data to capture a window of activity that could be used for subsequent biomechanical analysis.

### 4.6.6 Strengths, Limitations, and Future Research

This study is the first to describe the development of a human activity recognition model that (a) used training data collected from people who have knee osteoarthritis, (b) includes activities that are recommended by medical guidelines to monitor improvement in this population and (c) can identify not only activities but also phases of activities useful for biomechanical analysis. The training data for our human activity recognition model did not include a diverse range of participants, such as those who are severely disabled, obese, or have substantially different patterns of movement related to pain such as a 'step-to' gait pattern when using stairs. Future studies should include these diverse patient presentations to allow greater generalisability of a model.

Validation of this IMU-based human activity recognition model is required using data collected in conditions outside a laboratory environment (e.g. clinic or home). With further development, this model could be used in a workflow for analysing data collected when a patient was unobserved in order to segment data for subsequent biomechanical analysis using IMU fusion algorithms or machine learning predictions for knee joint kinematics (angles and speed) and kinetics (force or moments).

While the accuracy of our models was high for the first and second levels of classification, the accuracy for the third level of classification was substantially less. Subsequent studies should explore how to optimise model accuracy over multiple levels of classification.

Most of the published human activity recognition models have limited capacity to be used in unobserved conditions as people perform many diverse activities other than those on which the model was trained. The current model may therefore produce a significant number of false positive classifications in unobserved conditions. Future human activity recognition model development trained on data collected in laboratory conditions should also be validated for use in less controlled environments such as a clinic or in a person's home. The investigation of approaches that combine CNN with other machine learning approaches (e.g. CNN long short-term memory) is recommended as these approaches may further improve classification accuracy for activities that share similar features. Further investigation is warranted to explore the best number, location, and combination of sensors in the population with knee osteoarthritis.

## 4.7 Conclusions

Our results provide a proof-of-concept that data collected from people with knee osteoarthritis can be used to train human activity recognition models to classify clinically relevant activities, and phases of those activities that could be used for the purpose of monitoring an improvement due to treatment. This is the first study to develop a human activity recognition model from data collected from people who have knee osteoarthritis to classify clinically relevant activities and phases of activities in this population. The model accuracy was 85% at the first level of classification, 89% to 97% at the second level of classification and 60% to 67% at the third level of classification. The performance of our models compares well to other studies that classified different activities using the same validation approach (LOSOCV).

As we have demonstrated that these activities can be classified in the population, it may be possible to develop a human activity recognition system that can objectively measure the number of times a person who has knee osteoarthritis

performs an activity, as well as segment data to allow biomechanical analysis of these activities for data collected at home or at work.

## 4.8 Study Details

### 4.8.1 Author Contributions

Conceptualisation, J.-S.T., A.S., P.K., P.O., A.C.; data curation, J.-S.T., T.B., P.D.; formal analysis, J.-S.T., B.K.B., P.K., A.C., A.S.; investigation, J.-S.T., B.K.B., T.B., P.K., J.P.C., A.S., P.O., A.C.; methodology, J.-S.T., B.K.B., P.K., A.S., P.O., A.C.; project administration, J.-S.T., T.B., A.C.; resources, P.K., A.C.; supervision, J.P.C., P.K., A.S., P.O., A.C.; validation, J.-S.T., B.K.B.; visualisation, J.-S.T., B.K.B.; writing—original draft, J.-S.T.; writing—review and editing, J.-S.T., B.K.B., T.B., P.D., P.K., J.P.C., A.S., P.O., A.C. All authors have read and agreed to the published version of the manuscript.

### 4.8.2 Funding

An Australian Government Research Training Program Scholarship was received by the lead author to support his capacity to undertake this research.

### 4.8.3 Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Curtin University Human Research Ethics Committee (HRE2017-0695).

### 4.8.4 Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

### 4.8.5 Acknowledgments

We acknowledge Sawitchaya Tippaya for her assistance in data preparation.

### 4.8.6 Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## 4.9 Appendices

### Appendix 4-1

**Walking**

Stance

- Event marker 1: Initial heel contact of stance limb
- Event marker 2: Toe off of stance limb

Swing

- Event marker 1: Toe off of stance limb
- Event marker 2: Initial heel contact of stance limb

**Transitioning to and from a chair**

Sit-to-stand

- Event marker 1: From initiation of pelvis lift off (anterior/posterior movement of pelvic markers)
- Event marker 2: Maximum height of pelvis markers in standing

Stand-to-sit

- Event marker 1: Pelvis moving back down (anterior/posterior movement of pelvic markers)
- Event marker 2: Pelvis touch-down (anterior/posterior movement of pelvic markers ceases)

**Negotiating stairs**

Stance

- Event marker 1: Initial contact of stepping limb with next step
- Event marker 2: Toe off of stepping limb from the step

Swing

- Event marker 1: Toe off of swing limb from the step
- Event marker 2: Initial contact of swing limb with next step

**Table A1.** Classification Level 1—Chair, Stair, Walk.

| Learning Rate: 0.0001 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Convolution | | | | Transition | | Fully Connected | | | |
| **Type** | CONV2D | MAXPOOLING2D | CONV2D | MAXPOOLING2D | Flatten | Dense | Activation | Dropout | Dense | Activation |
| **Filter size** | 32 | | 64 | | | | | | | |
| **Kernel size** | 3 × 3 | 2 × 2 | 3 × 3 | 2 × 2 | | | | | | |
| **Function** | | | | | | | ReLu | | | Softmax |
| **Dropout size** | | | | | | | | 0.2 | | |
| **Neurons** | | | | | | 100 | | | 3 | |

**Table A2.** Classification Level 2—Stair Up/Down, Chair Sit/Stand.

| Learning Rate: Chair 0.0001, Stair 0.001 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Convolution | | | | | | | | | Transition | | Fully Connected | | | |
| **Type** | CONV2D | Activation | MAX POOL-ING2D | CONV2D | Activation | MAX POOL-ING2D | CONV2D | Activation | MAX POOL-ING2D | Flatten | Dense | Activation | Dropout | Dense | Activation |
| **Filter size** | 32 | | | 64 | | | 64 | | | | | | | | |
| **Kernel size** | 3 × 3 | | 2 × 2 | 3 × 3 | | 2 × 2 | 3 × 3 | | 2 × 2 | | | | | | |
| **Function** | | ReLu | | | ReLu | | | ReLu | | | | ReLu | | | Softmax |
| **Neurons** | | | | | | | | | | | 100 | | | 2 | |

**Table A3.** Classification Level 3—Walk Swing/Stance and Stair Up/Down Swing/Stance.

| Learning Rate: Walk/Stair Down 0.001, Stair Up 0.0001 | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Convolution | | | | | | | | Transition | | Fully Connected | | | |
| **Type** | CONV2D | Activation | MAXPOOLING2D | CONV2D | Activation | CONV2D | Activation | MAXPOOLING2D | Flatten | Dense | Activation | Dropout | Dense | Activation |
| **Filter size** | 32 | | | 64 | | 64 | | | | | | | | |
| **Kernel size** | 3 × 3 | | 2 × 2 | 3 × 3 | | 3 × 3 | | 2 × 2 | | | | | | |
| **Function** | | ReLu | | | ReLu | | ReLu | | | | ReLu | | | Softmax |
| **Neurons** | | | | | | | | | | 100 | | | 2 | |

## 4.10 Summary of Chapter 4

For clinicians working with people with knee osteoarthritis, the assessment of physical function across multiple clinically important activities is limited to direct observation during a clinical encounter. Therefore, clinicians lack objective means of monitoring performance of clinically important activities like walking, negotiating stairs and transitioning to and from a chair in free-living environments, such as when a patient is at home or work. While many studies had developed human activity recognition systems for healthy populations, no machine learning human activity recognition systems had been developed from and validated on data collected from people with knee osteoarthritis. We aimed to develop a human activity recognition system for people with knee osteoarthritis. We provide a proof-of-concept that a human activity recognition system can be developed from IMU data collected from people with knee osteoarthritis for clinically important activities and phases of activities. Human activity recognition models could be refined for use in monitoring performance of activities in free-living environments and act as the first part of a data handling pipeline to assist with segmenting and labelling data for subsequent biomechanical processing.

# Chapter 5

# Study 2b: Kinematic Prediction

## Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models

The results from the systematic review presented in Chapter 3 suggest there may be at least a proportion of the population with knee osteoarthritis where changing movement patterns is possible, and that change may be related to changes in clinical outcomes. One of the most investigated parameters was sagittal plane kinematics during walking. Within the systematic review one study demonstrated changes in sagittal plane angular kinematics during walking after Tai Chi (Zhu et al., 2016). There is also some indication in the literature outside of the systematic review that for people whose quadriceps strength increase, there is also a change in sagittal plane angular kinematics during walking (Davis et al., 2019). The systematic review did not find consistent group-level change in sagittal plane kinematics, but those findings do not suggest that individual-level change does not occur.

Notably, walking was the only clinically relevant functional activity that was investigated in the studies included in the systematic review. Other clinically relevant activities, such as negotiating stairs and transitioning to and from a chair, were not investigated despite guidelines recommending assessment all three of these activities.

Chapter 4 described the development of a human activity recognition system for multiple clinically relevant activities and phases of activities. This type of system has two potential uses (a) to monitor performance of clinically important activities in free-living environments, and (b) as the first part of a data handling pipeline that can segment and label the data for subsequent biomechanical analysis.

Here, Chapter 5 describes the development of an IMU-based kinematic prediction machine learning system for the prediction of sagittal plane kinematics during phases of multiple clinically important activities in people with knee osteoarthritis. This type of system could use segmented data output from the human activity recognition system presented in Chapter 4 as input to this subsequent step in a data handling pipeline.

**This chapter was published in the journal *Sensors*.**

**Tan, J.-S.**, Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., Kent, P., Smith, A., O'Sullivan, P., & Campbell, A. (2022). Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models. *Sensors*, *22*(2). https://doi.org/10.3390/s22020446

## 5.1 Abstract

Deep learning models developed to predict knee joint kinematics are usually trained on inertial measurement unit (IMU) data from healthy people and only for the activity of walking. Yet, people with knee osteoarthritis have difficulties with other activities and there are a lack of studies using IMU training data from this population. Our objective was to conduct a proof-of-concept study to determine the feasibility of using IMU training data from people with knee osteoarthritis performing multiple clinically important activities to predict knee joint sagittal plane kinematics using a deep learning approach. We trained a bidirectional long short-term memory model on IMU data from 17 participants with knee osteoarthritis to estimate knee joint flexion kinematics for phases of walking, transitioning to and from a chair, and negotiating stairs. We tested two models, a double-leg model (four IMUs) and a single-leg model (two IMUs). The single-leg model demonstrated less prediction error compared to the double-leg model. Across the different activity phases, RMSE (SD) ranged from 7.04° (2.6°) to 11.78° (6.04°), MAE (SD) from 5.99° (2.34°) to 10.37° (5.44°), and Pearson's r from 0.85 to 0.99 using leave-one-subject-out cross-validation. This study demonstrates the feasibility of using IMU training data from people who have knee osteoarthritis for the prediction of kinematics for multiple clinically relevant activities.

## 5.2 Introduction

People who have knee osteoarthritis commonly report pain and physical limitation performing functional activities such as walking, transitioning from a chair and negotiating stairs (Fukutani et al., 2016). During these activities they also use less sagittal plane range of movement (knee flexion) during particular phases of activities (e.g. stance phase of walking) compared to people who do not have osteoarthritis (Baliunas et al., 2002; Bouchouras et al., 2015; Hinman et al., 2002; McCarthy et al., 2013). Clinicians are interested in the relationship between specific kinematic measures and clinical outcomes in people with knee osteoarthritis (Tan, Tikoft, et al., 2021). For example, a person may have difficulty descending stairs because they do not use available knee flexion movement during the stance phase. Interventions such as exercise (Davis et al., 2019) and total knee replacement (Junsig Wang et al., 2019)

have demonstrated the ability to improve knee flexion angle during walking in people who have knee osteoarthritis. Clinical guidelines recommend that the performance of painful and limited activities are monitored over the course of treatment (Dobson et al., 2013). However, there are currently several limitations to clinicians being able to accurately quantify sagittal plane knee range of movement during functional activities in both clinical and free-living environments (e.g. patient's home or work, or during recreation).

Clinicians are unable to routinely access gold-standard optoelectronic motion analysis systems (e.g. Vicon) due to cost and space requirements. Smartphone camera-based technology is more accessible to clinicians and has demonstrated validity and reliability for measuring sagittal plane knee angles (Milanese et al., 2014). Both optoelectronic and smartphone camera-based systems require the patient to be observed within a fixed volume to record useful clinical information, precluding their use in a free-living environment. Inertial measurement units (IMUs) are a wearable sensor technology that is emerging as an alternative for biomechanical analysis, allowing a patient to move freely in clinical and free-living environments. Multiple scoping reviews have described the potential role of IMUs for the assessment of people with knee osteoarthritis (Cudejko et al., 2021; Kobsar et al., 2020) and following knee replacement surgery (Small et al., 2019). These reviews highlight the need for further investigation of IMU systems that can be used for monitoring biomechanics of patients in free-living environments.

There is a substantial volume of research validating IMUs for estimating kinematics in laboratory environments (Binnie et al., 2021; Rast & Labruyère, 2020; van der Straaten et al., 2018), although two barriers exist for widespread clinical adoption. In uncontrolled environments such as in a clinic or in free-living environments, the presence of metallic equipment (e.g. chairs or railings) and devices such as mobile phones and computers can interfere with the magnetometer data which can affect the reliability of fusion algorithm estimates (de Vries et al., 2009; Schall et al., 2016), making the data unusable (Schall et al., 2016). Although some fusion methods have been described which use only accelerometers and gyroscopes, they require IMU calibration prior to each use (Teufl et al., 2019). To overcome the

magnetometer problem and calibration requirements, machine learning (a form of artificial intelligence) approaches have been used to predict kinematics (e.g. knee joint flexion angle) from only the raw accelerometer and gyroscope data (Argent et al., 2019; Findlow et al., 2008). Although traditional machine learning requires the researcher to identify important features from the IMU data to train the model, a more contemporary approach is to use deep learning (a subfield of machine learning) that automatically detects features, minimising programming requirements (Hernandez et al., 2021; Mundt, Koeppe, David, Witter, et al., 2020; Rapp et al., 2021; Renani et al., 2021; Wouda et al., 2018).

There are a small number of studies where deep learning models have been trained to predict knee joint angular kinematics for walking from IMU training data collected mostly from healthy people (Hernandez et al., 2021; Mundt, Koeppe, David, Witter, et al., 2020; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Renani et al., 2021; Wouda et al., 2018). However, people with knee osteoarthritis experience significant difficulty with functional activities other than walking, such as negotiating stairs and transitioning to and from a chair. There is only one reported study using IMU data collected from participants who have knee osteoarthritis to train a deep learning model to predict sagittal plane knee kinematics, which was only for the activity of walking (Renani et al., 2021). No study has yet developed a deep learning model to predict knee joint kinematics for multiple, clinically important activities using IMU data collected from people with knee osteoarthritis.

The aim of this study was to demonstrate a proof-of-concept for the feasibility of using IMU training data collected from people who have knee osteoarthritis performing three clinically relevant functional activities: walking, negotiating stairs, and transitioning to/from a chair, to train a deep learning model to predict knee joint flexion angles. The second aim was to determine if a single-leg model (two sensors on one leg) or double-leg model (two sensors on both legs) was more accurate.

## 5.3 Materials and Methods

### 5.3.1 Study Design

This study was a retrospective, prognostic study using continuous IMU data, collected from people with knee osteoarthritis who performed multiple clinically relevant activities, to predict knee joint sagittal plane kinematics.

### 5.3.2 Participants

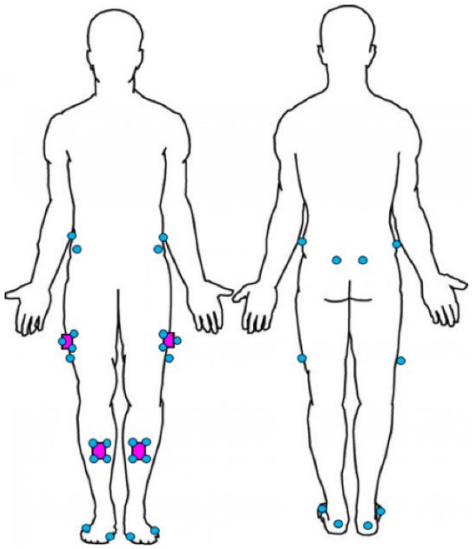
Participants in this study were part of a broader investigation into the use of IMUs in people with knee osteoarthritis (Binnie et al., 2021; Tan, Beheshti, et al., 2021). Seventeen participants with knee osteoarthritis were recruited from local physiotherapists, GP practices and local community centres. This number of participants mirrors other studies (Findlow et al., 2008; Stetter et al., 2020; Stetter et al., 2019; Wouda et al., 2018) and was thought to be sufficient to test the feasibility of this proof-of-concept study. Participants were included if they met the clinical diagnostic criteria for knee osteoarthritis (National Institute for Health & Care Excellence, 2014), had ≥3 months of pain, ≥4/10 pain on most days, and moderate activity limitation (single item on the Function, Daily Living sub-scale of the Knee injury and Osteoarthritis Outcome Score) (Roos & Lohmander, 2003). To minimise the effects of soft tissue artefact during motion capture that can introduce 'noise' into the data, we excluded participants with a body mass index (BMI) > 35 kg/m² and those who had a BMI >30 kg/m² with a waist-to-hip ratio (WHR) of ≤0.85 for women and ≤0.95 for men (those with greater soft tissue around the lower limbs). Participants were excluded if they had previous lower limb arthroplasty or mobility impairments due to other medical conditions (e.g. cognitive impairment, recent trauma, or neurological disorders). The study was approved by the Human Research Ethics Committee of Curtin University (HRE2017-0738).

### 5.3.3 Data Collection

Participants were initially screened for eligibility over the phone and subsequently attended a university motion analysis laboratory. After providing written informed consent, height and weight data were collected. IMUs and retroreflective markers were placed on the participants in a standardised manner by an experienced

musculoskeletal physiotherapist in the locations described in **Figure 5-1**. Participants performed 5 repetitions of knee flexion/extension as a warm-up on each knee. A standardised battery of functional activities was then performed that included 4 trials of stand-to-sit, 4 trials of sit-to-stand, 3 trials of 3-stair ascent, 3 trials of 3-stair descent, and 3 trials of a 5-metre self-paced walk. Participants rested for 30 s between trials and 60 s between activities. IMUs were removed after completion of the battery of functional activities and raw data were offloaded.

**Figure 5-1. IMU (purple) and Vicon marker (blue) placement.**



Bilateral Retroreflective marker locations:

- Anterior superior iliac spine
- Posterior superior iliac spine
- Superior iliac crest
- Lateral thigh between greater trochanter and lateral epicondyle 1-2cm below IMU
- 1st and 5th metatarsal head
- Calcaneal tuberosity
- 3 point cluster over thigh IMUs
- 4 point cluster over tibia IMUs

Bilateral IMU locations:

- Lateral thigh 2/3 distal to greater trochanter
- Anteromedial tibia 1/2 between tibial tuberosity and talocrural joint line

### 5.3.4 Instrumentation

Four IMUs (v6 research sensors, DorsaVi, Melbourne, Australia) sampling at 100 Hz (accelerometer 8G, gyroscope 2000 degrees/second) were attached to the lower limbs with double-sided hypoallergenic tape. The IMUs' dimensions were 4.8 × 2.9 × 1 centimetres, and they weighed 17 grams. Three-dimensional motion analysis was recorded with an 18 camera Vicon (Oxford Metrics Inc., Oxford, UK) sampling at 250 Hz. The relatively small reconstruction errors of <1 millimetre have resulted in the Vicon being considered the gold-standard motion analysis system (Ehara et al., 1995; Richards, 1999). Twenty-eight retroreflective markers were placed on the participant's pelvis and lower limbs using a cluster-based approach in alignment with International Society of Biomechanics recommendations (Wu et al., 2002). For this purpose, marker clusters were affixed to the IMUs (**Figure 5-1**), and anatomical markers were placed at the locations outlined in **Figure 5-1**. Additional markers were

applied to relevant joint centres for a static calibration trial, then removed (Besier et al., 2003), and the flexion/extension warm up trials were used to define the functional axis (Binnie et al., 2022). The sensor system was synchronised with the Vicon prior to being attached to the participant. IMUs in the same orientation were placed in a wooden box with retroreflective markers attached to the outside. The box was then rotated >90° ten times and recorded as a single trial in Vicon Nexus software (Oxford Metrics Inc., Oxford, UK) to facilitate subsequent time-synchronisation of the IMU and Vicon systems.

### 5.3.5 Data Preparation

Vicon trials were reconstructed and modelled using Vicon Nexus software. Gaps in trajectories were noted through visual inspection. Cubic spline interpolation was used to fill gaps of ≤20 frames (0.08 s), and if gaps were larger than this they were discarded. Kinematic trajectories were then filtered using a low-pass Butterworth filter with a 6 Hz cut-off frequency as determined by residual analysis. Vicon data were down-sampled from 250 to 100 Hz to allow time synchronisation with the IMU sensors.

We used the raw triaxial accelerometer and gyroscope data from 4 IMUs that were output as individual timestamped files using the IMU proprietary software (MDMv6 Manager v6.883, DorsaVi). Reconstructed Vicon data and the filtered raw orientation data from each IMU were time synchronised by the use of normalised cross-correlation using a customised LabVIEW program (National Instruments, Austin, TX, USA). The event markers were automatically detected by the LabVIEW program. Events for phases of walking and stair trials were heel contact and toe-off for swing and stance phases. Sit-to-stand and stand-to-sit events were anterior and posterior movement of the pelvis. Start and end times from the raw IMU data were exported for each phase of activity for the raw IMU and reconstructed Vicon data, which were used as inputs into the model. All trials were visually inspected to validate the automated synchronisation and event markers.

We input the affected leg, side of interest (leg being predicted), activity, direction of stair climbing and phase of activity as categorical variables into the model. In this study, we also investigate the interdependency between both legs by training

the model with two different structures of input data: double-leg, and single-leg. The double-leg model consisted of 38 input variables (24 accelerometer/gyroscope from 4 IMUs and 14 categorical variables), whereas the single-leg model included 27 input variables (12 accelerometer/gyroscope from 2 IMUs, side, 14 categorical variables). There were a total of 6 inputs from each IMU, with the accelerometer and gyroscope providing an input for each orthogonal (XYZ) axis (resulting in 24 input for the double-leg model, and 12 inputs for the single-leg model).

### 5.3.6 Deep Learning Model Development

The target prediction variable was the knee flexion joint angle at each time step obtained from the Vicon motion capture for multiple activities from the raw IMU accelerometer and gyroscope data.

One deep learning approach known as long short-term memory (LSTM) is suitable to handle discrepancies between steps in time-series data, where each trial differs in length (Mundt, Koeppe, Bamer, et al., 2020). LSTM also requires less pre-processing compared to other deep learning approaches, such as convolutional neural networks (CNNs), and is more suitable for real-time applications (Mundt et al., 2021). Recently a N-layer feed-forward neural network (FFNN) demonstrated superior results for kinematic prediction of the lower limb compared to a recurrent neural network known as LSTM (Mundt, Koeppe, David, Witter, et al., 2020). FFNN generally uses all the data points to make the prediction, whereas LSTM only uses past data points, resulting in its reduced accuracy for the first few data points. We chose to use a further evolution of LSTM and FFNN known as bidirectional LSTM (BiLSTM), which has both recurrent and feed-forward characteristics because it transverses the input data twice, using both past and future data for predictions (Renani et al., 2021) to improve accuracy compared to LSTM (Siami-Namini et al., 2019). BiLSTM has been successfully implemented for predicting knee joint kinematics during walking for people who have knee osteoarthritis or previous knee replacement (Renani et al., 2021).

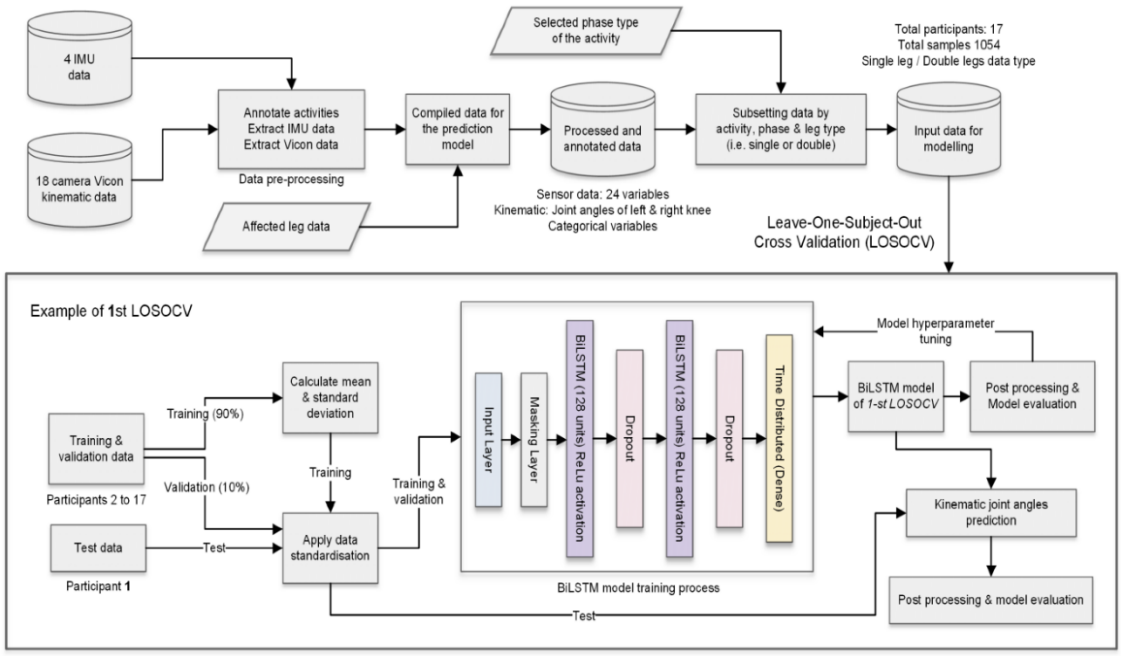
### 5.3.7 Model

A previous study using IMU data to train a deep learning prediction model reported that the number of IMUs can affect kinematic prediction error (Hendry et

al., 2021). To explore the effect of using additional IMUs, we developed two models: double-leg and single-leg. The double-leg model uses data from 4 IMUs from both legs as input to predict the knee joint angle of the leg of interest, whereas the single-leg model uses data from 2 IMUs from the leg of interest as input.

The model architecture was based on a stacked BiLSTM model. As BiLSTM requires input data for each sequence to have the same length, sequences were padded to the maximum sequence length of the activity phase. A single masking layer was used as the first layer to ignore all padded values. Then, two separate BiLSTM hidden layers with 128 units and a rectified linear unit activation function extracted the features from the sequences. BiLSTM was set to output a value for each time step in the input data, resulting in returning the sequence. A dropout layer was added after each BiLSTM hidden layer to randomly drop some units together with their connections from the network to reduce overfitting. The dropout rate was set to 0.2 and learning rate 0.0001. Finally, a separate fully connected time distributed output layer with linear activation was used to return the estimated joint angle (one for a single-leg input type and two for the double-leg input type). The proposed BiLSTM kinematic prediction model architecture is illustrated in **Figure 5-2**.

**Figure 5-2. Data preparation and model architecture of the proposed BiLSTM kinematic prediction models.**



The model is trained using the adaptive momentum (Adam) optimisation algorithm (Kingma & Ba, 2014). The final model parameters were selected after the hyperparameter tuning process assessing the loss function and model metrics. In this study, we retained the same hyperparameters across all model training processes to investigate the influence of input data variation to the prediction model. The data processing, machine model and experimental results were developed and implemented using Python 3 with the libraries Pandas, Numpy, Scipy, Scikit learn, Keras and Tensorflow.

### 5.3.8 Validation and Data Standardisation

As a clinician needs to know the average accuracy of a predictive model for each new patient, we used a leave-one-subject-out cross-validation (LOSOCV) method, which is most appropriate as it accounts for between-participant variability (Gholamiangonabadi et al., 2020). The LOSOCV method sequentially trained the model on the data from all participants except for one, which was left out and used for testing. This procedure looped through the total number of participants, resulting in 17 kinematic prediction models. The dataset was separated into training, validation and test datasets. For each validation fold, data from 16 participants were separated into training (90%) and validation (10%) sets randomly based on the unique samples. Training data were used to optimise the model parameters, whereas validation data were used as the unseen data during the model training process to fine tune the parameters such as validation loss, batch size and learning rate. Finally, the model was tested on all samples for the left-out participant's data.

For each activity, we calculated the average root mean square error (RMSE), normalised RMSE (nRMSE – RMSE / peak to peak amplitude) (Ren et al., 2008), mean absolute error (MAE) and Pearson correlation coefficient (r) between the Vicon reference and predictions for time-series data. Correlation coefficients were averaged across participants using Fisher's z transformation (Corey et al., 1998). The strength of the correlation was categorised as excellent ($r > 0.9$), strong ($0.67 < r \leq 0.9$), moderate ($0.35 < r \leq 0.67$) and weak ($r \leq 0.35$) based on similar studies (Stetter et al., 2020). In addition, we calculated the RMSE for the average maximum (peakRMSE) and minimum (minRMSE) knee flexion angles.

Each input variable and target variable were standardised for scale and distribution separately in each loop. Mean and standard deviation were computed only on the training data across all trials and all time points for each variable to prevent data leakage when applying pre-processing statistics. Validation and test data were standardised based on the corresponding mean and calculated from the training data used in each loop. A three-dimensional input shape is required for the BiLSTM model (N_samples, N_timesteps, N_features); therefore, input data are reshaped prior to being passed to the model. The number of samples collected from participants used to train the models is shown in **Table 5-1**.

**Table 5-1. Number of samples for each activity.**

| Phase of Activity | Samples (Participants) |
|---|---|
| Sit-to-stand | 61 (15) |
| Stand-to-sit | 61 (15) |
| Walk swing | 245 (17) |
| Walk stance | 244 (17) |
| Stair up swing | 130 (15) |
| Stair up stance | 87 (15) |
| Stair down swing | 83 (15) |
| Stair down stance | 44 (15) |
| Total | 955 (17) |

## 5.4 Results

The participant characteristics are shown in **Table 5-2**. The accuracy of our two models is presented in **Table 5-3**. Examples of representative model prediction for each activity phase compared to the Vicon reference standard (based on RMSE) are presented in **Figure 5-3.** Overall, the difference between the double-leg and the single-leg model was small, with an RMSE difference ranging from 0.11° to 1.96° and MAE from 0.01° to 1.46° for time-series predictions.

**Table 5-2. Characteristics of participants.**

| Characteristics | Mean (SD) |
|---|---|
| Age (years) | 66.2 (8.7) |
| Male (%) | 59% |
| Weight (kg) | 80.3 (15.9) |
| Height (cm) | 173 (8.8) |
| BMI (kg/m$^2$) | 26.6 (15.9) |
| KOOS function | 68.4 (12.6) |

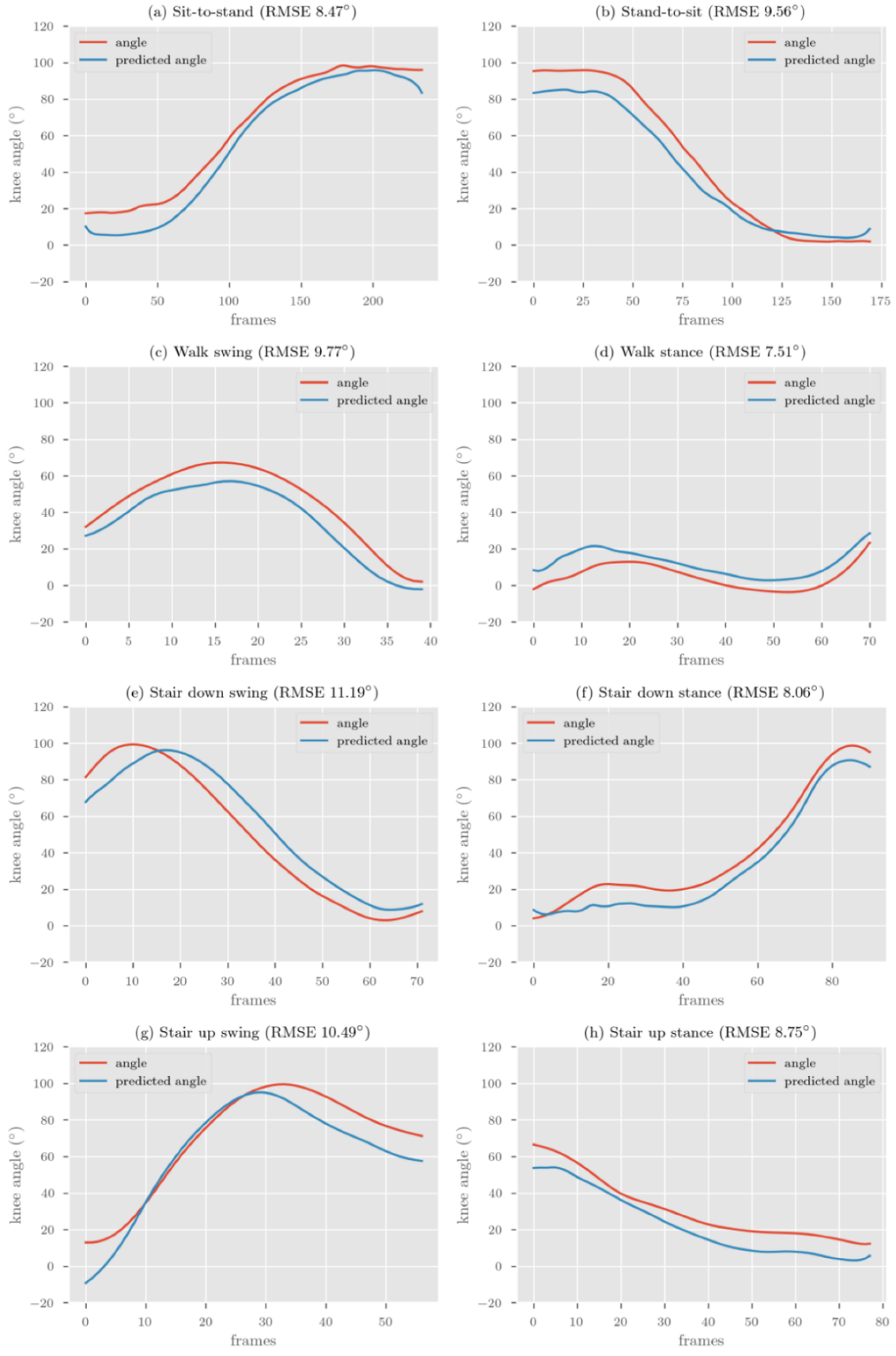
BMI = body mass index, cm = centimetres, kg = kilograms, KOOS = Knee injury and Osteoarthritis Outcome Scale, m = metres, SD = standard deviation.

**Table 5-3. Knee flexion angle prediction error for time-series, peak and minimum estimates for each activity.**

| Single-Leg Prediction Model | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Outcome | | Sit-to-stand | Stand-to-sit | Walk | | Stair Down | | Stair Up | |
| | | | | Swing | Stance | Swing | Stance | Swing | Stance |
| Time-Series | RMSE (°)(SD) | 8.24 (3.02) | 9.3 (2.99) | 9.7 (3.86) | 7.04 (2.60) | 11.78 (6.04) | 8.22 (2.80) | 10.41 (5.11) | 8.99 (3.70) |
| | nRMSE (%)(SD) | 9.79 (3.71) | 10.86 (3.78) | 17.66 (9.05) | 36.33 (14.39) | 14.06 (7.90) | 22.91 (9.99) | 15.06 (8.70) | 19.14 (10.00) |
| | MAE (°)(SD) | 7.12 (2.87) | 7.96 (2.60) | 8.46 (3.45) | 5.99 (2.34) | 10.37 (5.44) | 7.00 (2.55) | 9.06 (4.54) | 8.06 (3.64) |
| | r | 0.99 | 0.99 | 0.98 | 0.85 | 0.99 | 0.96 | 0.98 | 0.98 |
| Peak | RMSE (°)(SD) | 6.46 (2.48) | 6.89 (4.28) | 9.75 (6.21) | 10.31 (5.42) | 9.72 (3.72) | 21.38 (12.29) | 9.78 (6.65) | 11.73 (6.39) |
| Minimum | RMSE (°)(SD) | 6.92 (4.57) | 7.71 (5.77) | 7.35 (3.72) | 6.21 (2.99) | 8.07 (5.73) | 6.07 (4.69) | 10.33 (5.00) | 8.04 (5.76) |
| **Double-Leg Prediction Model** | | | | | | | | | |
| Outcome | | Sit-to-stand | Stand-to-sit | Walk | | Stair Down | | Stair Up | |
| | | | | Swing | Stance | Swing | Stance | Swing | Stance |
| Time-Series | RMSE (°)(SD) | 7.27 (1.72) | 8.10 (2.29) | 9.81 (3.98) | 8.19 (2.69) | 12.85 (5.63) | 10.19 (3.19) | 10.17 (4.63) | 9.61 (3.59) |
| | nRMSE (%)(SD) | 8.68 (2.58) | 9.45 (2.89) | 17.78 (8.68) | 43.33 (16.55) | 15.70 (7.45) | 32.93 (23.18) | 15.14 (8.29) | 19.90 (8.50) |
| | MAE (°)(SD) | 6.03 (1.69) | 6.72 (2.11) | 8.47 (3.52) | 6.92 (2.39) | 11.09 (5.07) | 8.47 (3.02) | 8.81 (4.25) | 8.36 (3.40) |
| | r | 0.99 | 0.99 | 0.97 | 0.74 | 0.98 | 0.92 | 0.98 | 0.96 |
| Peak | RMSE (°)(SD) | 5.09 (2.97) | 6.44 (4.23) | 9.23 (5.65) | 10.29 (6.51) | 10.73 (5.39) | 24.33 (10.70) | 10.01 (8.22) | 13.28 (8.18) |
| Minimum | RMSE (°)(SD) | 6.49 (4.55) | 6.15 (4.13) | 8.76 (4.31) | 6.60 (2.37) | 11.21 (8.60) | 8.99 (3.79) | 10.36 (5.02) | 7.79 (5.36) |

° = degrees of movement, MAE = mean absolute error, r = Pearson correlation coefficient, RMSE = root mean squared error, nRMSE = normalised RMSE, SD = standard deviation.

**Figure 5-3.** Representative single-leg BiLSTM model prediction compared to Vicon reference for each activity phase.



The single-leg model demonstrated the smallest RMSE/MAE across activities (five of eight activities) for time-series predictions compared to the double-leg model. The double-leg model demonstrated the smallest RMSE/MAE for predicting simultaneous double-leg activities (sit-to-stand and stand-to-sit) compared with activities that require reciprocal movements of both legs (walking or stairs).

Correlations between the reference Vicon system and the deep learning model were excellent ($r > 0.9$) for both the single-leg and double-leg models for all activities except for the stance phase of walking. The strongest correlation coefficient was for sit-to-stand, stand-to-sit and the swing phase of ascending stairs ($r = 0.99$).

The peakRMSE and minRMSE for each activity was almost always lower than the time-series RMSE for large range activities (sit-to-stand, stand-to-sit and swing phases). For small range activities (stance phases), the minRMSE was always lower than the time-series RMSE, whereas the peakRMSE was always higher.

## 5.5 Discussion

The aim of this study was to establish a proof-of concept for the feasibility of using IMU data collected from people who have knee osteoarthritis for development of a deep learning model to predict sagittal plane knee joint angles for multiple clinically relevant activities. We developed a BiLSTM kinematic prediction model on IMU training data that included walking, negotiating stairs, and transitioning to/from a chair for people who have knee osteoarthritis. The prediction error (RMSE/MAE) between the reference Vicon system and kinematic predictions was lowest for the stance phase for walking and going down stairs, and highest for the swing phase for going down and up stairs. Although, as a proportion of the range used during each activity phase, the sit-to-stand and stand-to-sit had the lowest prediction error, the stance phase for walking and going down stairs had the highest prediction error (nRMSE). For time-series data, the shape of the predicted curve had a consistently excellent correlation ($r > 0.9$) to the Vicon system across activities.

The second aim was to develop two types of models using training data from (a) two IMUs on one leg (single-leg) and (b) four IMUs on two legs (double-leg). The single-leg model demonstrated more frequent smaller errors and had excellent

correlations (r > 0.9) than the double-leg model across activities that require reciprocal, asymmetrical lower limb movement, such as walking and negotiating stairs; however, the difference in error between models was small. For activities that require bilateral simultaneous movement (sit-to-stand/stand-to-sit), the double leg model demonstrated smaller error.

### 5.5.1 Comparison to Previous Literature

In comparison to other deep learning models were that developed to predict knee joint kinematics just for walking, our BiLSTM model for multiple activities demonstrated prediction errors and correlations within the range of previous studies (RMSE 0.97° to 12.1°, r 0.94 to 0.99) (Hernandez et al., 2021; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Renani et al., 2021; Wouda et al., 2018).

Only one other study has used deep learning to predict knee kinematics in people who have knee osteoarthritis for the activity of walking (Renani et al., 2021). Our model has the benefit of being more broadly applicable for real world use, when combined with human activity recognition, because of the inclusion of multiple clinically important activities for people who have knee osteoarthritis. The model by (Renani et al., 2021) demonstrated small average RMSE 2.9° (SD 1.1°) for time-series prediction of knee flexion/extension during walking using a BiLSTM model. Training data in that study was from four IMUs placed on the pelvis, thigh, shank and foot. In comparison, our BiLSTM model trained on data from only a thigh and shank IMU demonstrated substantially higher average RMSE during walking phases (single-leg model – stance 7.04° (SD 2.6°), swing 9.7° (SD 3.86°)). Their results may have demonstrated lower error because of the higher number of samples (n = 3943) compared to our study (n = 955), the additional IMUs placed on the pelvis and foot, the inclusion in our training data of activities other than walking, or the difference in validation approach.

For validation of a IMU prediction model to be meaningful to a clinician, it has been suggested that the average level of error for each new person should be reported (Gholamiangonabadi et al., 2020), which is a strength of the LOSOCV method compared to the other validation methods (e.g. k-fold cross validation). Our results are similar to those of previous studies that use LOSOCV (Wouda et al.,

2018), rather than studies that use other validation approaches (Hernandez et al., 2021; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Renani et al., 2021), and studies that use real IMU data compared to those that use simulated IMU data (see section 5.5.3.3.). For example, (Wouda et al., 2018) validated a LSTM model using IMU training data collected from a healthy population for time-series prediction of knee flexion during running. They used LOSOCV and reported an average RMSE of 12.1° (SD 1.5°). In comparison, our model achieved lower average RMSE for walking swing and walk stance of 9.7° (SD 3.8°) and 7.0° (SD 2.6°). Although our model demonstrated lower average error, there was higher variability, which may be the result of using training data that included multiple activities rather than the single activity of walking. Our model also demonstrated good ability to predict the shape of the kinematic curve with excellent correlations ($r > 0.9$) for time-series prediction of all but one activity, comparing well to the models by Wouda et al. (2018) ($r = 0.94$) and Renani et al. (2021) ($r = 0.99$).

Using raw accelerometer and gyroscope data for training deep learning prediction models appears a promising tool to aid clinical decision making for clinicians managing people with movement disorders, such as knee osteoarthritis, as it mitigates the requirement for the magnetometer, which is prone to interference, especially in free-living environments where the magnetic field is not uniform (Weygers et al., 2020). However, deep learning approaches using real IMU data for the prediction of knee kinematics have not yet reached the consistent low error achieved by Kalman filter-based approaches that report RMSE as low as 1° for multiple clinically relevant activities (Teufl et al., 2019) or 5.04° using the proprietary software for the IMUs described in this study (Binnie et al., 2021). Various clinical, data handling and machine learning architecture considerations may help to reduce prediction error in future studies.

### 5.5.2 Clinical Considerations for Kinematic Prediction Models

Development of various machine learning models has the potential to have a significant impact for clinical populations, such as for people who have knee osteoarthritis. However, the majority of these studies have not described the clinical implications of such models; therefore, this section discusses clinical considerations

for future development of machine and deep learning models for prediction of joint kinematics.

#### 5.5.2.1 Variability of Movement in Clinical Populations

Although our model demonstrated a very small error and a high correlation for some participants, this was not the case across all participants. It is well established that people who have musculoskeletal or neurological health conditions move differently than healthy populations (Astephen et al., 2008; Junsig Wang et al., 2019; Zanardi et al., 2021). Across people who have knee osteoarthritis, there is diversity in movement patterns during functional activities related to disease severity (Astephen et al., 2008). Because of this heterogeneity of movement patterns across different conditions and even within a single diagnosis such as like knee osteoarthritis, it is important that models are trained and tested on the intended population for use. For example, (Renani et al., 2020) trained a CNN to predict spatiotemporal kinematics of the lower limb for people with knee osteoarthritis and after total knee replacement. Their model demonstrated consistently higher prediction error and variability across 12 spatiotemporal gait parameters for people that have knee osteoarthritis compared to people who had total knee replacement. Other studies have reported that the accuracy of human activity recognition models derived on data from healthy populations has had substantially reduced test accuracy in people who have health conditions, such as Parkinsonism (Albert et al., 2012; Lonini et al., 2016). It is currently unknown if the test accuracy of kinematic prediction models differs across populations. Given that people with knee osteoarthritis move differently and more variably than healthy people, the ability to generalise the kinematic prediction model accuracy between those populations should not be assumed. Future studies should consider testing prediction models on participants with health conditions of interest who demonstrate a range of movement impairments, and pain and disability levels.

#### 5.5.2.2 Selecting Clinically Important Activities and Movement Parameters

The clinically relevant use for predicting sagittal plane knee joint angles is to monitor biomechanics during functional activities in free-living environments and in-clinic to aid clinical decision making. Specifically, particular phases of activities

(e.g. stance phase of ascending stairs) are of interest to clinicians because they are targets for rehabilitation.

Prior to this study, machine learning models for predicting knee joint kinematics that could potentially be useful for people with knee osteoarthritis have only been trained and tested on walking data (Findlow et al., 2008; Hernandez et al., 2021; Mundt, Koeppe, David, Witter, et al., 2020; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Renani et al., 2021; Renani et al., 2020). Unlike those studies where a kinematic prediction model was developed for walking, our model is the first to be trained and tested on a range of clinically relevant activities for a specific clinical population. Although walking is the most frequently performed activity of the lower limbs, we selected three activities (walking, negotiating stairs and transitioning to/from a chair) that are recommended as part of a clinical physical assessment in medical guidelines for knee osteoarthritis (Dobson et al., 2013). To improve clinical utility of machine learning prediction using IMU data, future studies should investigate kinematic prediction models for a broader range of clinically important activities.

There are a broad range of kinematic and kinetic movement parameters that are of interest to clinicians and researchers for people with knee osteoarthritis (Tan, Tikoft, et al., 2021). Therefore, monitoring sagittal plane knee joint angles is only one movement parameter that could be recorded for clinically relevant activities in free-living environments. Other movement parameters are also of interest because of their relationship with structural progression of knee osteoarthritis. Knee adduction moment, for example, is associated with the progression of medial compartment knee osteoarthritis (Chehab et al., 2014; Miyazaki et al., 2002). There is early work investigating spatiotemporal kinematics (Renani et al., 2020), predicting knee moments and forces using deep learning approaches such as LSTM, CNN and ANN for the purposes of field monitoring (Mundt, Thomsen, et al., 2020; Stetter et al., 2020; Stetter et al., 2019). To improve clinical utility of IMU machine/deep learning prediction models using IMU data, future studies should investigate integrating (Kobsar et al., 2020) human activity recognition (Tan, Beheshti, et al., 2021) with both

kinematic and kinetic prediction models (see section 5.5.3.1) for a broad range of clinically relevant activities that include but are not limited to the activity of walking.

#### 5.5.2.3 Reducing the Burden for Clinicians

Some studies use up to 17 IMUs across the whole body to train deep learning models for kinematic prediction of the lower limbs (Renani et al., 2020; Wouda et al., 2018). It is generally thought that having additional IMUs results in improved accuracy and reduced error for machine learning predictions using IMU data for human activity recognition (Hendry et al., 2020; Lee et al., 2020). However, having to use additional IMUs can be burdensome for clinicians. Our findings are similar to (Hendry et al., 2021), who investigated kinematic prediction for the hip and lumbar spine using IMUs to train a deep learning model for ballet dancers. They reported that their kinematic prediction model trained on only two IMUs placed on the lower limbs demonstrated less error (7.0°) than models that included additional training data from IMUs placed on the spine (7.8°). A novel finding in our study was that the prediction error with using data from only two IMUs on a single leg was less than that using four IMUs on two legs, which may be because of the asymmetrical and diverse nature of movement patterns that exist in people with knee osteoarthritis. Future studies should aim to determine the minimum number of IMUs required for specific conditions and activities, to reduce clinician burden.

### 5.5.3 Considerations for Future Data Handling and Machine Learning Models

#### 5.5.3.1 Developing Data Handling Pipelines

Previously published kinematic (Findlow et al., 2008; Hernandez et al., 2021; Mundt, Koeppe, David, Witter, et al., 2020; Mundt, Thomsen, et al., 2020; Rapp et al., 2021; Renani et al., 2021; Renani et al., 2020; Wouda et al., 2018) and kinetic (Mundt, Koeppe, David, Witter, et al., 2020; Stetter et al., 2020; Stetter et al., 2019) prediction models are currently only useful in conditions in which the wearer of the IMUs is observed, such as in a clinical environment. This is because in free-living environments people wearing IMUs will perform other activities in addition to walking (e.g. transitioning to/from a chair and negotiating stairs). Therefore, biomechanical prediction models have limited use in free-living environments without

additional data processing that can automate the identification and labelling the long, continuous streams of data that are produced by IMUs.

Our approach was to train the kinematic prediction model on labelled data that could potentially be output from a human activity recognition deep learning algorithm as part of a data handling pipeline. We previously established a proof-of-concept about the development of a human activity recognition model (Tan, Beheshti, et al., 2021) that can segment data into the phases of clinically important activities described in this study, which could be the first component of a data-handling pipeline.

However, it is currently unknown which method of data segmentation is most useful, minimally burdensome for clinicians and computationally efficient. We selected phases of activities because clinicians are typically interested in data from phases, rather than the whole gait cycle (see section 5.5.2.2). Data in other studies has been segmented in a variety of ways including continuous walking (Hernandez et al., 2021), three gait cycles (Findlow et al., 2008), or single gait cycles (Rapp et al., 2021; Renani et al., 2021). The higher-order data segmentation in those studies may prove to be clinically useful for use in a human activity recognition model that includes other activities (e.g. going up stairs or sit-to-stand), and integration with a gait event detection algorithm (Fadillioglu et al., 2020).

#### 5.5.3.2 Single vs. Multiple Models

We developed a single kinematic prediction model to include training data from multiple activities, which provides more generalisability and precludes the need to model every activity (Stetter et al., 2019). However, there is uncertainty about the superiority of universal single models for prediction of kinematics across multiple activities compared to multiple models that predict only specific activities. (Stetter et al., 2019) reported the development of an ANN to predict knee joint forces for 16 sports specific activities (e.g. walking, running, jumping, and cutting). They noted the possibility that their model had higher error compared to the study by (Wouda et al., 2018) was because of their use of a single model for the multiple activities (Stetter et al., 2019). Contrary to this, our single model for multiple activities had lower RMSE

(Wouda et al., 2018) and stronger correlations (Findlow et al., 2008) than other approaches that only used training data from a single activity.

Furthermore, our double-leg model performed better than the single-leg model for bilateral simultaneous activities of sit-to-stand and stand-to-sit. (Stetter et al., 2019) demonstrated a similar effect where there a single-leg model had higher error for two leg activities (jump take-off and two leg jump landing) compared to single-leg activities.

These results together may indicate that, in future studies, activity-specific models should be directly compared to models trained to predict kinematics for multiple activities, and that models trained on both legs may impact the results of asynchronous movement such as walking.

#### 5.5.3.3 Augmented and Simulated Data

One challenge of developing generalisable kinematic prediction models is the collection of a sufficient number of samples from a representative cohort of participants, a process which is burdensome. This challenge is highlighted by the large RMSE/MAE for the stair down stance, where there was the least number of training data for the model. One solution becoming increasing popular is to include augmented or simulated training data (Mundt, Koeppe, David, Witter, et al., 2020; Renani et al., 2021). Data augmentation involves manipulating the data by offsetting or warping the time or magnitude (Renani et al., 2021), while simulation involves developing synthetic IMUs from other biomechanical data sources (Mundt, Koeppe, David, Witter, et al., 2020). It has been demonstrated that by augmenting or simulating data results in a 27% to 45% improvement in RMSE, with improvements in knee flexion RMSE between 1.4° to 5.22° (Dorschky et al., 2020; Renani et al., 2021). These approaches using simulated and augmented data may provide additional benefit in models trained on data collected from clinical populations, such as people with knee osteoarthritis. In addition, using augmented data may help reduce the impact of misplacement of sensors by either clinicians or patients.

#### 5.5.3.4 Deep Learning Architecture

We used BiLSTM following on from the work of (Renani et al., 2021). BiLSTM is proposed to improve prediction accuracy because it transverses the input data twice, using both past and future data points, compared to traditional LSTM (Siami-Namini et al., 2019). Future studies should investigate the performance of BiLSTM compared to traditional LSTM and other deep learning approaches for kinematic prediction. Although we used BiLSTM, there is some indication that combining multiple deep learning architectures, such as CNN with LSTM (ConvLSTM), can improve prediction accuracy for IMU data (Ascioglu & Senol, 2020). Hernandez et al. (2021) demonstrated that this combined deep learning approach using ConvLSTM can provide good results for knee flexion time-series predictions (MAE 3° (SD 1.15°), r = 0.99) using a nested k-fold validation with a 70% training, 15% validation and 15% test approach (Hernandez et al., 2021). Researchers must further investigate the balance between predictive accuracy and the requirement for pre-processing of data. Mundt et al. (2021) tested the predictive accuracy of a CNN, LSTM and multilayer perceptron network for lower limb kinematics and kinetics. They demonstrated superior accuracy with a CNN for prediction of kinematics, although the pre-processing requirements are high for this type of model compared to LSTM, which may be more suited to real-time applications.

### 5.5.4 Limitations

Because this study was a proof-of-concept investigation, there are a number of limitations. We included only 17 participants, did not have a representative number of female participants, and excluded people with high BMI. These factors may limit the generalisability of our model for the broader population with knee osteoarthritis. Further, there was an unbalanced dataset with a significantly different number of trials across activities, which may have affected the results. This study included clinically relevant activities described in clinical guidelines for people who have knee osteoarthritis (Dobson et al., 2013). However, there are additional activities that people perform daily that were not included. A single model for predicting kinematics for multiple activities was used in this study, which may have affected the prediction error, and it is unclear if a universal model is feasible for all activities a person may

perform. Further investigation is required to determine the comparative accuracy of a single model for multiple activities versus individual models for each type of clinically relevant activity (or phase of activities) in clinical populations, such as people who have knee osteoarthritis.

## 5.6 Conclusions

This proof-of-concept study demonstrates that using IMU training data collected from people who have knee osteoarthritis to predict sagittal plane knee joint kinematics during multiple clinically important activities using a deep learning model is feasible. Our novel BiLSTM model demonstrated that using training data from as few as two IMUs placed on one leg performs with less error for most activities than with additional training data from IMUs on both legs. To be of clinical value, the model presented in this study could be combined with a human activity recognition system to monitor response to treatment in people with knee osteoarthritis.

## 5.7 Study Details

### 5.7.1 Author Contributions

Conceptualisation, J.-S.T., A.S., P.K., P.O., A.C.; data curation, J.-S.T., T.B., K.N., P.D.; formal analysis, J.-S.T., S.T., P.K., A.C., A.S.; investigation, J.-S.T., S.T., T.B., P.D., K.N., P.K., J.P.C., A.S., P.O., A.C.; methodology, J.-S.T., S.T., P.D., K.N., P.K., A.S., P.O., A.C.; project administration, J.-S.T., T.B., A.C.; resources, P.K., A.C.; supervision, J.P.C., P.K., A.S., P.O., A.C.; validation, J.-S.T., S.T.; visualisation, J.-S.T., S.T.; writing—original draft, J.-S.T.; writing—review and editing, J.-S.T., S.T., T.B., P.D., K.N., P.K., J.P.C., A.S., P.O., A.C. All authors have read and agreed to the published version of the manuscript.

### 5.7.2 Funding

An Australian Government Research Training Program Scholarship was received by the lead author to support his capacity to undertake this research.

### 5.7.3 Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Curtin University Human Research Ethics Committee (HRE2017-0695).

### 5.7.4 Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

### 5.7.5 Data Availability Statement

The data presented in this study may be available on request from the corresponding author.

### 5.7.6 Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## 5.8 Summary of Chapter 5

Clinicians do not have access to laboratory-based motion analysis systems in clinical practice to help inform clinical reasoning. While IMU systems exist for biomechanical assessment in clinical practice, they are prone to error in environments where it is not possible to control electromagnetic interference, which impacts the reliability of fusion algorithms. We aimed to develop a sagittal plane angular kinematic machine learning prediction model from IMU data collected from people with knee osteoarthritis for multiple clinically important activities that was robust to electromagnetic interference, preventing the need for calibration. We developed a single model for predicting knee flexion angle during phases of walking, ascending, and descending stairs, and transitioning to and from a chair. We found that the single-leg model had less prediction error for walking and negotiating stairs compared to the double-leg model that was better for transitioning to and from a chair. This kinematic prediction model has the potential to be the second part of a data handling pipeline that uses data previously segmented from a human activity recognition prediction model.

# Chapter 6

# Study 2c: Kinetic Prediction

## Deep Learning for Predicting Moments and Forces from Wearable Sensors in People with Knee Osteoarthritis

Kinetic movement parameters such as knee adduction and flexion moment are associated with structural progression and symptoms associated with knee osteoarthritis (Chehab et al., 2014; D'Souza et al., 2022; Wilson et al., 2021). However, routine clinical practice does not currently involve monitoring changes in kinetic parameters. Chapter 4 provided a foundation for segmenting and labelling walking data using human activity recognition for stance and swing phases. Chapter 5 provided a proof-of-concept that movement parameters can be predicted from IMU training data collected from people with knee osteoarthritis with comparable accuracy to studies in healthy people when a leave-one-subject-out cross-validation is used.

While the systematic review presented in Chapter 3 suggests that there is an infrequent relationship between change in movement patterns and clinical outcomes, this conclusion was mostly based on group-level data. There is some indication in people with knee osteoarthritis, that strong relationships exist for individualised changes in movement patterns and clinical outcomes (Kobsar & Ferber, 2018). Previously published biomechanical prediction models for people with knee osteoarthritis have trained models on group data, as is standard practice, but for individualised prediction, additional data from the test participant may help reduce prediction error.

The results from Chapter 5 also suggested that a single-leg model (using training data from four IMUs placed across two legs) provided superior accuracy for predicting sagittal plane angular kinematics for ambulatory activities compared to a

double-leg model (two IMUs placed on one leg). No studies had explored if double-leg or single-leg models are more accurate for kinetic parameters during walking.

Therefore, the aims of this study of predicting kinetics were to explore (a) the effect on the prediction error of using leave-one-subject-out cross-validation compared to participant-specific validation, and (b) if a single-leg model (two IMUs on one leg) or a double-leg model (four IMUs on two legs) was more accurate.

**A manuscript from this chapter has been submitted to a journal and is currently under review.**

**Tan, J.-S.**, Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., Smith, A., O'Sullivan, P., Campbell, A., & Kent, P. Deep learning for predicting moments and forces from wearable sensors in people with knee osteoarthritis.

## 6.1 Abstract

Clinicians are unable to assess abnormal knee loading of people with knee osteoarthritis in clinical practice or in free-living environments because of limitations accessing technology that is inexpensive, robust to electromagnetic interference and does not require calibration. The objective of this study was to train a deep learning model to predict knee joint moments and forces from accelerometer and gyroscope training data collected from people with knee osteoarthritis to explore: (a) the effect on prediction error for different validation approaches (leave-one-subject-out cross-validation (LOSOCV) and participant-specific validation), and (b) the effect of training the models on data from one or two legs. We trained bidirectional long short-term memory models on data from 16 participants with knee osteoarthritis. The range of the results for moments were: normalised RMSE (nRMSE) (SD) = 22% (5%) to 43% (15%), r = 0.71 to 0.8, and for forces: nRMSE = 16% (4%) to 57% (23%), r = 0.78 to 0.86 for forces. Prediction error (nRMSE) for the participant-specific validation was 9% to 36% lower than LOSOCV and the double-leg models were 1% to 23% lower than the single-leg models. This study demonstrates the feasibility of using movement sensor training data to predict knee moments and forces in people with knee osteoarthritis, and that prediction error is influenced by personalising a model and training those models on data from two legs rather than one. We developed clinically relevant deep learning models to predict knee kinetic parameters from wearable sensor data that could help monitor abnormal knee loading in free-living environments.

## 6.2 Introduction

There is a substantial body of literature describing the clinical implications of abnormal knee joint loading in people with knee osteoarthritis (D'Souza et al., 2022; Tan, Tikoft, et al., 2021). After a diagnosis of knee osteoarthritis has been established, abnormal knee joint moments (e.g. increased knee adduction or flexion moments) (Chehab et al., 2014; D'Souza et al., 2022) and forces (e.g. medial contact force) (Wilson et al., 2021) during the stance phase of walking have been implicated as risk factors for structural progression of knee osteoarthritis and eventual knee

replacement (Hatfield et al., 2015). Therefore, there are a number of interventions that clinicians can provide that aim to reduce abnormal knee loading such as exercise (Tan, Tikoft, et al., 2021), orthotics (Radzimski et al., 2012), braces (Moyer et al., 2015) and surgical approaches (Fantini Pagani et al., 2020).

Clinicians are unable to access traditional motion analysis systems assess knee joint loading because they are expensive, require dedicated space and require technical expertise. Therefore, clinicians are unable assess joint moments or forces in clinic or in patient's free-living environment (home or work) to help inform clinical decision making around selection of interventions. Inertial measurement units (IMUs) are one type of wearable movement sensor that could be used during consultations for a person with knee osteoarthritis.

Hardware in IMUs typically include an accelerometer, gyroscope, and magnetometer which do not directly measure kinetic parameters. Therefore, the earliest approaches for estimating forces and joint moments using IMUs required kinematic modelling and calibration (Karatsidis et al., 2019; Koning et al., 2015; Konrath et al., 2019). However, electromagnetic interference can affect the magnetometer, and therefore kinematic modelling, making IMUs less reliable in uncontrolled conditions (e.g. free-living environment). One approach that does not require kinematic modelling nor calibration is to use machine learning to predict kinetic parameters from raw accelerometer and gyroscope data.

Machine learning approaches, such as recurrent neural networks (long-short term memory (LSTM)) (Mundt, Koeppe, David, Bamer, et al., 2020; Mundt, Thomsen, et al., 2020) and feed-forward neural networks (Mundt, Koeppe, David, Witter, et al., 2020; Stetter et al., 2020; Stetter et al., 2019; Wang et al., 2020; Wouda et al., 2018), have been used to predict knee joint forces (Stetter et al., 2019) and knee joint moments (Stetter et al., 2020; Wang et al., 2020) from IMU data during the stance phase of gait of healthy people without knee osteoarthritis. There are two studies that have used IMU data collected from people with knee osteoarthritis to predict knee adduction moment (He et al., 2019; Wang et al., 2020), although models for predicting other moments or knee joint forces have not been investigated in this population.

Individualised assessment for people with knee osteoarthritis is recommended in clinical guidelines. There is indication in the literature that participant-specific machine learning approaches result in improved accuracy compared to population-based models trained on data from healthy participants (Ahamed et al., 2019; Findlow et al., 2008; Rodríguez-Martín et al., 2017). However, the magnitude of this improvement has yet to be tested in the population with knee osteoarthritis.

A recent study has demonstrated improved prediction accuracy for models trained on data from two legs to predict sagittal plane kinematics during walking compared to models trained on one leg (Tan et al., 2022). Although this has not yet been tested for kinetic models.

The aim of this study was to develop a deep learning model to predict knee moments and forces using training data collected from people with knee osteoarthritis to explore (a) the effect on the prediction error using leave-one-subject-out cross-validation compared to participant-specific validation, and (b) if a single-leg model (two IMUs on one leg) or a double-leg model (four IMUs on two legs) was more accurate. This study is exploratory and was not designed as a head-to-head comparison with other prediction model algorithmic approaches.

## 6.3 Methods and Materials

We conducted an exploratory study using IMU data collected from people with knee osteoarthritis to train deep learning models to predict time-series and peak knee joint moments and forces for the stance phase of walking.

### 6.3.1 Participants

We recruited seventeen participants with knee osteoarthritis from local physiotherapists, general practitioners, and local community centres. Participants were included if they had a clinical diagnosis of knee osteoarthritis (National Institute for Health & Care Excellence, 2014), for more than three-months, with moderate (≥4/10) pain on most days and moderate activity limitation (single item on the Function, Daily Living sub-scale of the Knee injury and Osteoarthritis Outcome Score) (Roos & Lohmander, 2003). We excluded people with a body mass index (BMI) >35kg/m$^2$ and those who had a BMI >30kg/m$^2$ with a waist-to-hip ratio

(WHR) of ≤0.85 for women and ≤0.95 for men (those with greater soft tissue around the lower limbs) to minimise the potential for soft tissue artefact that can result in noise in the IMU signal. Participants were also excluded if they had mobility impairments unrelated to their knee osteoarthritis and if they had any cognitive impairments that would prevent them from being able to participate in data collection procedures. The institutional research ethics committee approved this study (HRE2017-0738).

### 6.3.2 Data Collection and Instrumentation

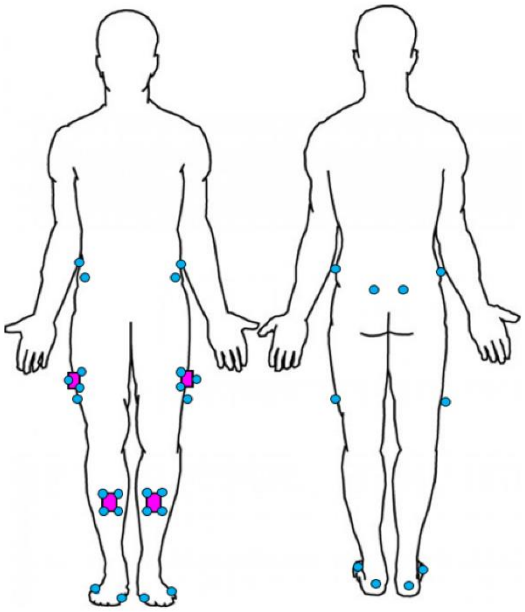
Data were collected in a motion analysis laboratory. Participant's height and weight were recorded using a stadiometer and calibrated electronic scales after providing informed written consent. Four IMUs (v6 research sensors, DorsaVi, Melbourne, Australia) sampling at 100Hz were then attached to the participant by an experienced musculoskeletal physiotherapist in a standardised manner according to the manufacturer's instructions using double sided hypoallergenic adhesive (**Figure 6-1**).

An 18-camera Vicon MX motion analysis system (Oxford metrics Inc., Oxford, UK) (sampling frequency 250Hz) and two AMTI force plates (AMTI, Watertown, USA) (sampling frequency 2000Hz) recorded data subsequently used to estimate the kinetic parameters used as the ground-truth (reference standard). This system has reconstruction errors of <1mm and is considered the gold-standard motion analysis system (Richards, 1999; Wu et al., 2002). Fourteen retro-reflective Vicon markers were placed on anatomical landmarks, and an additional fourteen markers over the four IMUs in accordance with a cluster based approach (Besier et al., 2003) that follows International Society of Biomechanics recommendations (Wu et al., 2002) (**Figure 6-1**). A static calibration trial was conducted for the Vicon system (Besier et al., 2003), with additional 10 retroreflective markers placed over the medial/lateral femoral epicondyles and malleoli, and on the calcaneum. Participants were asked to stand in their normal posture for five seconds to capture their neutral knee joint position, and additional markers were removed prior to motion capture. A functional axis approach was used to define the joint centre (Binnie et al., 2022) using flexion/extension trials prior to the participant performing the functional activities.

The IMU and Vicon systems were synchronised by placing the IMUs in a wooden box with retroreflective markers attached to the outside then rotating box >90° ten times and recorded as a single trial in Vicon Nexus software (Oxford Metrics Inc., Oxford, UK). Each participant's data was recorded in a deidentified manner.

The participants were then instructed to perform a standardised battery of functional activities that included: a warm up for five repetitions of knee flexion/extension, transitioning from a chair (five trials of sit-to-stand-to-sit), negotiating stairs (three trials of a three-stair ascent, three trials of a three-stair descent) and walking (three trials of a five metre self-paced walk) which were collected for another study (Tan, Beheshti, et al., 2021). For this study only data from the stance phase of walking trials were used (three to four stance phases per trial).

**Figure 6-1. IMU (purple) and Vicon marker (blue) placement.**



Bilateral Retroreflective marker locations:

- Anterior superior iliac spine
- Posterior superior iliac spine
- Superior iliac crest
- Lateral thigh between greater trochanter and lateral epicondyle 1-2cm below IMU
- 1st and 5th metatarsal head
- Calcaneal tuberosity
- 3 point cluster over thigh IMUs
- 4 point cluster over tibia IMUs

Bilateral IMU locations:

- Lateral thigh 2/3 distal to greater trochanter
- Anteromedial tibia 1/2 between tibial tuberosity and talocrural joint line

### 6.3.3 Data Processing

Data from each IMU were offloaded and output as timestamped files. The Vicon and force plate data were analysed post hoc in Vicon Nexus Software (Oxford Metrics Inc., Oxford, UK). Breaks in marker trajectories were visually inspected. We used cubic spline interpolation for breaks <20 frames (0.08 seconds) and for larger breaks the trial was discarded. A 6 Hz low pass digital filter (Butterworth) was applied, determined from a residual analysis. Inverse dynamics modelling was performed using a validated, reliable three-dimensional mathematical model (Besier et

al., 2003; Wu et al., 2002). Marker trajectories and ground reaction force data from the walking trials were output as timestamped files that were then used to calculate knee joint moments and forces.

Movement data were further processed using a custom LabVIEW program (National Instruments, Texas, USA). First, the Vicon data were down-sampled to 100 Hz to allow direct comparisons with the IMU data, and the two datasets were time-synchronised using normalised cross-correlation. Next, time points for the beginning and end of each stance phase were identified in the Vicon data. Finally, the program compiled and output the peak and time-series knee joint moments and forces for the Vicon data. Because the input data sequences were of different length, we time-normalised both sets of data to 101 data points (representing 0% to 100% of the stance phase). Joint moment and force amplitudes were normalised to body weight. Knee moments were expressed as external moments. Finally raw triaxial accelerometer and gyroscope data from the IMUs was organised into 24 columns for the double-leg model, and 12 columns for the single-leg model as sensor inputs into the machine learning model.

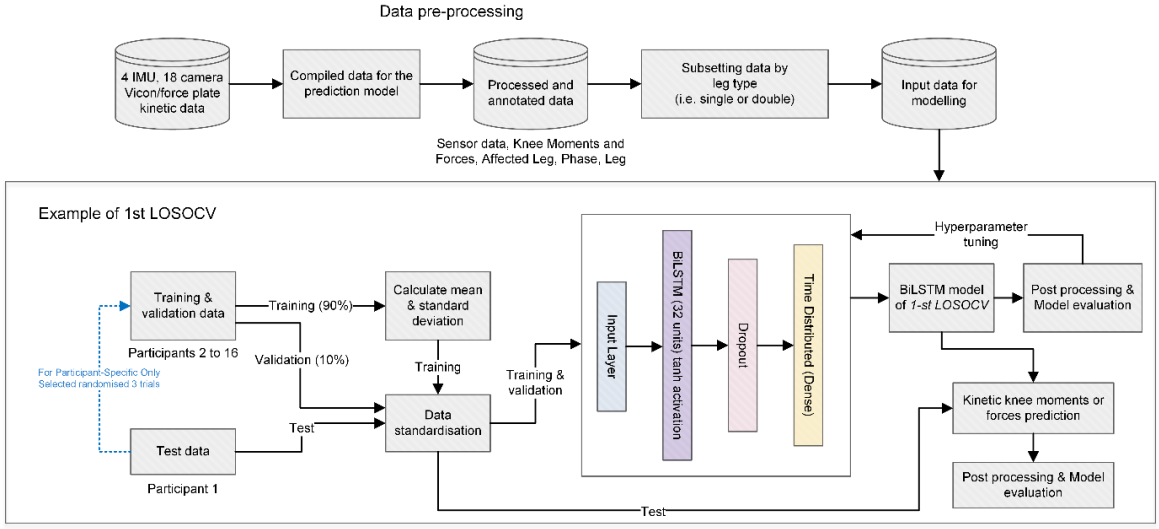
### 6.3.4 Development of the Machine Learning Model

The target prediction variables derived from Vicon and force plate data are were knee adduction and flexion moments, and knee compression and medial contact force. Separate models were developed for moments and forces. Input data included raw IMU accelerometer and gyroscope input data and categorical variables; affected leg, stance phase, and side. We selected a deep learning approach known as bidirectional long-short term memory (BiLSTM), previously used for predicting knee kinematics (Renani et al., 2021; Tan et al., 2022). BiLSTM allows for use of more data from which to make predictions as both past and future data points are used as input which make more accurate predictions for time-series data (Schache et al., 2008) compared with traditional LSTM (recurrent neural network) that only uses past data to make prediction or feed-forward neural networks that uses future data (Mundt, Koeppe, David, Witter, et al., 2020).

We used a single BiLSTM model. Features were extracted using a single BiLSTM hidden layer of 32 units and a hyperbolic tangent activation function. For

each time step, the BiLSTM was set to output a corresponding value. To prevent overfitting, we used a dropout layer (dropout rate 0.3) after BiLSTM layer. The final layer was fully connected (dense) time distributed output layer with linear activation that provided the predicted output variables. **Figure 6-2** depicts the BiLSTM kinetic prediction model workflow. An adaptive momentum (Adam) optimiser algorithm was used to train the model (Kingma & Ba, 2014). We optimised the hyperparameters by assessing the loss function and model metrics to determine the final model. Python 3 was used for data processing, development of the machine learning model and processing final results using Pandas, Numpy, Scipy, Scikit learn, Keras, and Tensorflow libraries.

**Figure 6-2. BiLSTM kinetic prediction model workflow.**




### 6.3.5 Statistical Analysis

Two validation methods were used to test the accuracy of the model. First, the leave-one-subject-out cross-validation (LOSOCV) approach trained the model using all data except for one participant's, which was used as hold out data for testing the model. Second, a participant-specific approach extended the LOSOCV method, by including all but three trials from the single participant hold out data to the training data, with the remaining three trials used as the test set.

For each LOSOCV fold, the data were randomised into 90% training (model optimisation), 10% validation (tuning parameters), then tested on the hold out data. For the participant-specific model, the additional training data were first added to the

larger dataset, prior to randomisation. Tuning the model parameters during the validation phase included adjusting the validation loss, batch size and learning rate.

As a measure of model performance, the average accuracy of time-series data predictions were calculated, including root mean square error (RMSE), normalised RMSE (nRMSE) (RMSE normalised to the range of the data) (Ren et al., 2008), mean absolute error (MAE) and Pearson correlation coefficient (r). Fisher's z transformation was used to reduce the bias of r value sampling distribution skew (Corey et al., 1998). Correlation coefficients were categorised as excellent ($r > 0.9$), strong ($0.67 < r \leq 0.9$), moderate ($0.35 < r \leq 0.67$) consistent with studies in this field (Stetter et al., 2019). The RMSE for the average maximum knee joint moments and forces were also calculated.

For each validation loop, the scale and distribution of the input and output variables were standardised. To prevent data leakage during pre-processing of statistics, the mean and standard deviation were calculated only on training data for each trial and timepoint. Subsequently the validation and test data were standardised based on the resultant mean from the training data for each loop. The input data into the BiLSTM model were reshaped into three-dimensional matrix with N_samples, N_timesteps and N_features.

## 6.4 Results

The demographic data and average of the peak of the moments and forces for the 16 participants are presented in **Table 6-1**. Because of errors with their force plate data, one participant was removed from the dataset. For each participant, there were three trials of three to four stance phases, resulting in 157 observations across the cohort. The prediction error for the kinetic prediction models are presented in **Table 6-2**. **Appendix 6-1** provides a representative visual example of the prediction error between the model and reference standard (based on nRMSE) for each kinetic movement parameter. The difference in nRMSE between LOSOCV and participant-specific validation are presented in **Table 6-3** and the difference in nRMSE between double-leg and single-leg models is presented in **Table 6-4**.

**Table 6-1. Participant Characteristics**

| Characteristics | All participants Mean (SD) |
|---|---|
| Age (years) | 65.4 (8.3) |
| Male (%) | 50% |
| Weight (kg) | 80.6 (16.5) |
| Height (cm) | 173.9 (8.5) |
| BMI (kg/m$^2$) | 26.5 (4.4) |
| KOOS physical function | 67 (13) |
| Peak | |
| Flexion moment (Nm/kg)* | 0.53 (0.29) |
| Adduction moment (Nm/kg)* | 0.58 (0.13) |
| Compression force (N/kg)* | 9.74 (0.79) |
| Medial force (N/kg)* | 0.74 (0.72) |

BMI = body mass index, cm = centimetres, kg = kilograms, KOOS = Knee injury and Osteoarthritis Outcome Scale, m = metres, SD = standard deviation.*mean peak to minimum difference

**Table 6-2. Knee joint moment and force prediction error for time-series, and peak estimates during the stance phase of walking.**

| | | RMSE (SD) | | nRMSE(%) (SD) | | MAE (SD) | | r | Peak RMSE (SD) | |
|---|---|---|---|---|---|---|---|---|---|---|
| Flex moment (Nm/kg) | | | | | | | | | | |
| LOSOCV | Single | 0.23 | (0.08) | 37 | (14) | 0.19 | (0.07) | 0.78 | 0.27 | (0.08) |
| | Double | 0.22 | (0.07) | 36 | (13) | 0.19 | (0.07) | 0.78 | 0.27 | (0.11) |
| PS | Single | 0.19 | (0.05) | 24 | (5) | 0.15 | (0.04) | 0.81 | 0.26 | (0.16) |
| | Double | 0.18 | (0.04) | 23 | (5) | 0.15 | (0.03) | 0.80 | 0.19 | (0.14) |
| Adduction moment (Nm/kg) | | | | | | | | | | |
| LOSOCV | Single | 0.20 | (0.06) | 43 | (15) | 0.17 | (0.06) | 0.75 | 0.23 | (0.11) |
| | Double | 0.18 | (0.06) | 37 | (13) | 0.16 | (0.06) | 0.79 | 0.21 | (0.09) |
| PS | Single | 0.14 | (0.03) | 24 | (5) | 0.12 | (0.03) | 0.71 | 0.12 | (0.06) |
| | Double | 0.13 | (0.03) | 22 | (5) | 0.11 | (0.02) | 0.81 | 0.13 | (0.06) |
| Compression force (N/kg) | | | | | | | | | | |
| LOSOCV | Single | 2.05 | (0.74) | 27 | (10) | 1.71 | (0.69) | 0.80 | 1.49 | (0.67) |
| | Double | 1.96 | (0.63) | 25 | (9) | 1.61 | (0.58) | 0.82 | 1.24 | (0.6) |
| PS | Single | 1.61 | (0.36) | 17 | (4) | 1.27 | (0.3) | 0.84 | 1.05 | (0.53) |
| | Double | 1.48 | (0.36) | 16 | (4) | 1.12 | (0.27) | 0.86 | 1.00 | (0.56) |
| Medial force (N/kg) | | | | | | | | | | |
| LOSOCV | Single | 0.52 | (0.16) | 57 | (23) | 0.45 | (0.14) | 0.78 | 0.41 | (0.19) |
| | Double | 0.45 | (0.13) | 34 | (10) | 0.39 | (0.12) | 0.82 | 0.37 | (0.19) |
| PS | Single | 0.34 | (0.07) | 21 | (4) | 0.28 | (0.06) | 0.84 | 0.26 | (0.14) |
| | Double | 0.31 | (0.06) | 19 | (4) | 0.25 | (0.05) | 0.86 | 0.23 | (0.16) |

SD = standard deviation, RMSE = root mean square error, nRMSE = normalised root mean square error, MAE = mean absolute error, r = Pearson's correlation coefficient, LOSOCV = leave-one-subject-out cross-validation, PS = participant-specific

**Table 6-3. Difference in nRMSE (%) between LOSOCV and participant-specific validation for single-leg and double-leg models (positive favours participant-specific)**

| | Single-leg model (%) | | | Double-leg model (%) | | |
|---|---|---|---|---|---|---|
| | LOSOCV | Participant-specific | *Difference* | LOSOCV | Participant-specific | *Difference* |
| Flexion moment | 37 | 24 | **13** | 36 | 23 | **13** |
| Adduction moment | 43 | 24 | **19** | 37 | 22 | **15** |
| Compression force | 27 | 17 | **10** | 25 | 16 | **9** |
| Medial force | 57 | 21 | **36** | 34 | 20 | **14** |

LOSOCV = leave-one-subject-out cross-validation

**Table 6-4. Difference in nRMSE (%) between double-leg models and single-leg models for LOSOCV and participant specific validation (positive favours double-leg models)**

| | LOSOCV (%) | | | Participant-specific (%) | | |
|---|---|---|---|---|---|---|
| | Single-leg | Double-leg | *Difference* | Single-leg | Double-leg | *Difference* |
| Flexion moment | 37 | 36 | **1** | 24 | 23 | **1** |
| Adduction moment | 43 | 37 | **6** | 24 | 22 | **2** |
| Compression force | 27 | 25 | **2** | 17 | 16 | **1** |
| Medial force | 57 | 34 | **23** | 21 | 20 | **1** |

LOSOCV = leave-one-subject-out cross-validation

The participant-specific cross-validation approach demonstrated less prediction error than the LOSOCV approach for both single-leg and double-leg models for all kinetic parameters. The double-leg prediction model had less error than the single-leg prediction model for all kinetic parameters. The difference in the nRMSE between the double-leg and single-leg prediction models ranged substantially (1% to 23%) for the LOSOCV approach but did not differ considerably for participant-specific model (1% to 2%).

The model with the lowest error was the double-leg participant-specific model. Comparing the results for that model, compression force had the lowest nRMSE (16%), followed by medial force (20%), adduction moment (22%) and flexion moment (23%), with strong correlations (>0.8) with the Vicon-force plate ground-truth.

## 6.5 Discussion

We aimed to develop a machine learning model to predict knee joint moments and forces for the stance phase of walking using IMU training data collected from people who have knee osteoarthritis. We explored the effect of training the models on additional participant-specific data and the difference in prediction error for models trained on data from two legs (double-leg model) compared to data from one leg (single-leg model). The results indicate that participant-specific models demonstrated less prediction error compared to LOSOCV, and that double-leg models have lower prediction error compared to the single-leg models.

Previous studies developed machine learning models to predict knee adduction moment in people with knee osteoarthritis (He et al., 2019; Wang et al., 2020). Building on those studies, our models also predict time-series flexion moment, compression and medial forces and their peaks. Together, these studies provide the groundwork for future IMU-machine learning kinetic prediction models for people with knee osteoarthritis that could be used for screening of biomechanical risk factors and monitoring change in movement patterns in response to interventions but are not intended to be an input to the clinical diagnosis of knee osteoarthritis. This study also provides novel insights into how the number of IMUs, and personalising prediction models can affect model prediction error.

Across studies that have used training data from people with knee osteoarthritis there are differences in normalisation procedures for knee joint moments, statistical reporting, and validation approaches, which limits some comparisons. He et al. (2019) reported the prediction error for time-series knee adduction moment using a feed-forward neural network (RMSE = 0.36 Nm/kg*m and r = 0.91) using an unspecified validation approach. Wang et al. (2020) reported prediction error (RMSE <0.004 Nm/kg*m, $R^2$ = 0.947) using a decision tree-based approach. The prediction error for their alternative artificial neural network model was reported differently (MAE = 0.004 Nm/kg*m or <20% of average knee adduction moment, with an $R^2$ = 0.956) using leave-one-out cross-validation (Wang et al., 2020). Our most comparable model (single-leg, LOSOCV) had a prediction error of RMSE 0.20 Nm/kg, r = 0.75. We

normalised moments and forces to body weight alone, while those studies normalised moments to body weight multiplied by height in metres (He et al., 2019) and centimetres (Wang et al., 2020). While our model demonstrated lower correlation compared to He et al. (2019), the differences in normalisation limit meaningful comparisons between RMSE/MAE results.

Further limiting comparison is the inconsistency in cross-validation approaches used (Janidarmian et al., 2017). We selected LOSOCV because it would provide an estimate of the expected error for each new patient a clinician sees in clinical practice. It is not possible to directly compare results of Wang et al. (2020) and He et al. (2019) as they did not use LOSOCV. While Wang et al. (2020) reported using a leave-one-out cross-validation approach for their artificial neural network, the validation approaches differ. When using leave-one-out cross-validation, data is shuffled across participants prior to input into the model, meaning that the model is not evaluated on data from a single participant during each fold, underestimating the prediction error for a new participant's data. There is a clear need for consistent validation methods and reporting standards to facilitate comparison between IMU-machine learning studies.

A more direct comparison can be made between the current study and two studies using IMU data from 13 healthy participants to train a deep learning model to predict knee joint moments (Stetter et al., 2020) and forces (Stetter et al., 2019) during the stance phase of walking. In those studies, feed-forward neural network models were trained on data from IMUs placed over the anterior thigh and shank and validated using a LOSOCV approach. They reported prediction error for knee adduction moment (RMSE = 0.18 Nm/kg, nRMSE = 22.3%, r = 0.71) and knee flexion moment (RMSE 0.26 = Nm/kg, nRMSE = 18.4%, r = 0.72), compression force (nRMSE = 14.2% and r = 0.87) and medial force (nRMSE = 27.7% and r = 0.6). Our single-leg model validated using LOSOCV produced lower RMSE, higher nRMSE, and stronger correlations than the two feed-forward neural network models (Stetter et al., 2020; Stetter et al., 2019). Because of the complexity of clinical, statistical and machine learning factors involved in building a biomechanical prediction model there are a range of factors that can result in variable prediction

accuracy across studies. A range of factors could explain the difference in results between studies including differing movement patterns between people with knee osteoarthritis and those without, the number and type of ambulatory activities, age of participants, sensor placement, ground-truth kinematic modelling, number of training samples and the type of machine learning model architecture (e.g. BiLSTM vs feed-forward neural network). Further investigation to explore the individual effects of clinical, statistical and machine learning factors is clearly warranted.

Machine learning models are most accurate when the variability in the data is low. Therefore, accuracy of IMU-machine learning prediction models can be affected because of the diversity in movement patterns across people with knee osteoarthritis. For example, one way to reduce the variability maximally is to train the model only on data from the test participant, rather than from a population-based model (Albert et al., 2012). However, collecting enough training samples is burdensome in a clinical environment. So, one way to accommodate for between participant variability is to refine the population-based model by the addition of participant-specific training data to the population-based training dataset. We demonstrated that adding participant-specific training data to the model improves nRMSE by 9% to 15% for the double-leg models and 10% to 36% for the single-leg models. This finding is consistent with other studies that have demonstrated that participant-specific IMU-machine learning prediction models outperform generic models for the prediction of lower limb kinematics (Findlow et al., 2008) and running pattern classification (Ahamed et al., 2019) in healthy people, and improves detection of freezing of gait in people with Parkinson's disease (Rodríguez-Martín et al., 2017). It is not possible to have the ground-truth (e.g. Vicon and force plate) data in a routine clinical encounter, systems would need to be created that allow patient-specific training data to be uploaded, matched to similar IMU data from a database that does have paired ground-truth data, and used to train a cloud-based real-time dynamic model building system for kinetic predictions.

We previously found that single-leg models provide lower sagittal plane kinematic prediction errors for ambulatory activities compared to a double-leg models (Tan et al., 2022). Unlike those findings, the current study demonstrates lower

prediction error for double-leg compared to single-leg models. One possibility for this difference is that there may be a stronger relationship between kinetic parameters and bilateral leg movement than for kinematics. Another reason may be that there is a stronger relationship between movement patterns of both legs in the frontal plane than the sagittal plane. This is supported in the current study where when using the LOSOCV approach, medial force predictions were nRMSE 23% different, and adduction moment were nRMSE 6% different, while flexion moment and compression force were only 1% to 2% different between double-leg and single-leg models. There is consistency across this study and previous studies (Stetter et al., 2020; Stetter et al., 2019) that prediction error for frontal plane movement parameters is higher than other movement planes. Further research should confirm if prediction error can be improved for frontal plane movement predictions using IMU combinations that consider both legs.

### 6.5.1 Limitations

The population sampled was of 'normal' BMI which may have limited the generalisability of the results. Future research could investigate the current models are generalisable to populations with higher BMI. We selected kinetic prediction outcomes for the stance phase of walking as this measure is related to the structural progression of medial knee osteoarthritis. Therefore, it is unknown if the current models would be suitable for other functional activities, nor for data in free-living environments. Although consistent with prior research, the current sample size may limit generalisability of the model. Future studies could use larger samples that may include simulated or augmented data to overcome data collection requirements of additional participants (Mundt, Koeppe, David, Witter, et al., 2020; Renani et al., 2021).

### 6.5.2 Conclusion

We have demonstrated the development of a deep learning approach for predicting knee joint moments and forces from raw accelerometer and gyroscope data in people with knee osteoarthritis. Training the BiLSTM model on data from four IMUs placed on two legs resulted in less prediction error that two IMUs on one leg. Participant-specific models had less prediction error than LOSOCV. IMU-machine

learning models could eventually be used in clinical practice to assess and monitor abnormal knee loading in people with knee osteoarthritis in clinical and free-living environments.

## 6.6 Study Details

### 6.6.1 CRediT Authorship Contribution Statement

Jay-Shian Tan: Writing – original draft, Data Acquisition, Methodology, Investigation, Formal analysis, Data curation, Conceptualisation, Funding acquisition.

Sawitchaya Tippaya: Writing – review & editing, Methodology, Investigation, Formal analysis, Data curation, Study design, Software, Visualisation.

Tara Binnie: Writing – review & editing, Data Acquisition, Investigation.

Paul Davey: Writing – review & editing, Data curation, Investigation, Software.

Kathryn Napier: Writing – review & editing, Data curation.

J.P. Caneiro: Writing – review & editing, Investigation.

Anne Smith: Writing – review & editing, Investigation, Formal analysis, Methodology, Conceptualisation.

Peter O'Sullivan: Writing – review & editing, Investigation, Conceptualisation.

Amity Campbell: Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualisation.

Peter Kent: Writing – review & editing, Methodology, Investigation, Formal analysis, Methodology, Conceptualisation, Supervision.

### 6.6.2 Declaration of Competing Interest

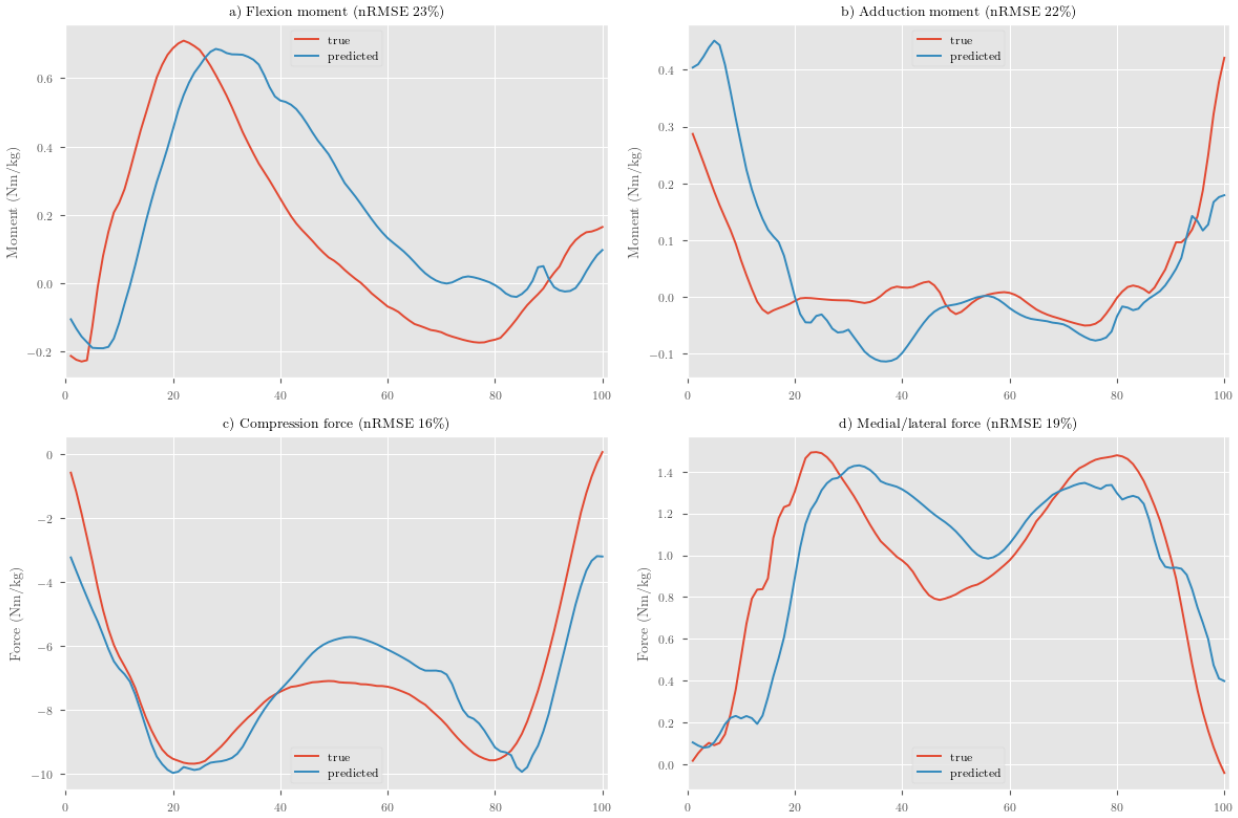
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### 6.6.3 Acknowledgements

An Australian Government Research Training Program Scholarship was received by the lead author to support his capacity to undertake this research.

## 6.7 Appendix

**Appendix 6-1 Representative double-leg BiLSTM participant-specific model prediction (blue) compared to reference standard (red) for the stance phase of gait.**



## 6.8 Summary of Chapter 6

Increased knee loading (e.g. knee adduction moment) is a risk factor for structural progression of knee osteoarthritis. Clinicians managing patients with knee osteoarthritis do not have access to technology for monitoring knee loading in free-living environments. Generalisable models that provide predictions about an individual whose data was not included in machine learning model training have been previously developed for people with knee osteoarthritis. The effect of adding additional data from the test participant to the generalisable model had not been tested for predicting moments and forces in people with knee osteoarthritis. We aimed to explore (a) the effect of adding training data from a test participant to the generalisable model, and (b) if a single-leg model or double-leg model was more accurate. Adding data from the test participant consistently reduced prediction error for all moments and forces. The double-leg prediction model demonstrated less prediction error than the single-leg model for the stance phase of walking, in contrast to the findings in Chapter 5. With further testing and implementation, such a moment and force prediction model could be used secondary to human activity recognition prediction model to provide clinicians the ability to monitor knee loading in patients with knee osteoarthritis in clinical practice and free-living environments.

# Chapter 7
# Discussion of Thesis

The aims of this thesis were to:

1. Systematically review cohort studies and randomised controlled trials to investigate how changes in knee joint movement parameters during functional activities relate to changes in activity limitation or pain after exercise intervention in people with knee osteoarthritis.
2. Investigate how wearable sensor technology could be used to monitor activity avoidance and altered movement patterns in people with knee osteoarthritis:
    a. Develop an IMU-based, human activity recognition system that can classify clinically relevant activities and phases of activities (walking, negotiating stairs, and transitioning to and from a chair) for people with knee osteoarthritis;
    b. Develop machine learning prediction models for knee joint sagittal plane angular kinematics for multiple clinically important activities; and
    c. Develop machine learning prediction models for knee joint moments and forces for the stance phase of walking.

This chapter discusses the main findings of this thesis through a clinical lens within the context of using machine learning approaches based on IMU data for monitoring physical function and movement patterns, and how these may relate to clinical outcomes of patient-reported activity limitation and pain. The strengths and limitations of this thesis are then described followed by a broad consideration for the integration of this technology in future clinical practice and research.

## Summary of Findings

### Systematic Review – Chapter 3

**Research question**

Is there a relationship between changes in movement patterns and clinical outcomes?

**What this research adds**

The systematic review found a relationship between change in movement parameters and change in clinical outcomes occurred 24% of the time using group-level data (from 20 studies), and 13% of the time using individual-level data (from two studies).

**Secondary findings**

- The most frequent movement parameters that had been investigated were:
  - First peak knee adduction moment (20 occasions);
  - Knee flexion moment (14 occasions);
  - Knee adduction moment impulse (8 occasions); and
  - Various measures of knee flexion angle (8 occasions).
- Gait-retraining consistently resulted in both changes in movement parameters and in clinical outcomes:
  - Measures of knee adduction moment were the movement parameter that changed most frequently in a predictable, directional manner.
  - After gait retraining, the prevalence of any movement parameter changing was 45% (5/11 times tested).
  - The expected direction of change in knee adduction moment (Simic et al., 2013) occurred 100% of the time.

**Comparison to previous literature**

This study built upon a prior systematic review that investigated the effect of exercise on changes in knee adduction moment during walking (Ferreira et al., 2015) that did not find evidence of a relationship across two included studies.

| This systematic review (Tan, Tikoft, et al., 2021) | (Ferreira et al., 2015) |
|---|---|
| Included studies | |
| • 22 cohort studies and randomised controlled trials | • Three randomised controlled trials |
| Quantification of relationship | |
| • Relationship quantified using within-group change or correlation analysis | • Relationship quantified using between-group difference |
| Movement parameters | |
| • Knee kinematics, moments, and muscle activity | • Only knee adduction moment |
| Activities included in search strategy | |
| • Investigated any functional activity | • Investigated only walking |

## Summary of Findings

### Human Activity Recognition – Chapter 4

**Research question**

What is the accuracy of a human activity recognition (HAR) system trained on IMU data from people with knee osteoarthritis to classify clinically relevant activities and phases?

**What this research adds**

This study is the first to report on the development and validation of a machine learning HAR system trained on IMU data from people with knee osteoarthritis. Classification accuracy of the deep neural network (CNN model) for a three-level classifier was 87% (activity), 89-97% (direction) and 60-67% (phase), and equivalent to the prediction accuracy for HAR systems trained on data from healthy people.

**Secondary findings**

- There was a reduction in accuracy between the first and third levels of classification, with an increase in accuracy between the first and second level of classification.

**Comparison to previous literature**

This study builds on earlier studies as previous HAR systems developed for healthy people have only classified activities, not the phases of activities, and those phases are useful for biomechanical analysis. Gholamiangonabadi et al. (2020) developed a CNN HAR system for 12 activities for healthy participants and reported a 69-79% accuracy. Also, for more granular levels of classification, that accuracy can reduce by 4-16% (Hendry et al., 2020; Whiteside et al., 2017).

| Human activity recognition model (Tan, Beheshti, et al., 2021) | (Gholamiangonabadi et al., 2020) |
|---|---|
| **Participants** | |
| • 18 with knee osteoarthritis | • 10 healthy participants from benchmarking dataset |
| **Activity** | |
| • Level 1 (chair, stairs, walking)<br>• Level 2 (sit down, stand up, stairs ascending, stairs descending)<br>• Level 3 (stance and swing of walking and stairs) | • Standing, sitting, walking, lying, stairs, waist bends forward, elevation of arms, crouching, cycling, jogging, running, jumping |
| **Number and location of IMUs** | |
| • Four IMUs (bilateral thigh and shank) | • Three IMUs (wrist, chest, ankle) |
| **Validation approach** | |
| • Leave-one-subject-out cross-validation | • Leave-one-subject-out cross-validation |

## Summary of Findings

### Kinematic Prediction – Chapter 5

**Research questions**

1. What is the prediction error of a machine learning prediction model for sagittal plane flexion angles, trained on IMU data from people with knee osteoarthritis?
2. What is the prediction error of a single-leg model compared to a double-leg model?

**What this research adds**

This study is the first to report on the development of a BiLSTM model using IMU data for the prediction of knee flexion angle during multiple clinically important activities for people with knee osteoarthritis.

Overall, the single-leg model (RMSE (SD) = 7.0° (2.6°) to 11.8° (6.0°), and r = 0.85 to 0.99) demonstrated lower prediction error than the double-leg model (RMSE (SD) = 7.3° (1.7°) to 12.9° (5.6°), and r = 0.74 to 0.99).

**Secondary findings**

- The single-leg model outperformed double-leg model for asymmetrical activities (such as walking and stairs).
- The double-leg model outperformed the single-leg model for symmetrical activities (sit-to-stand and stand-to-sit).

**Comparison to previous literature**

Only one other study has developed a BiLSTM model to predict knee flexion angle using IMU data collected from people with knee osteoarthritis. Renani et al. (2021) reported lower prediction error for their model (RMSE (SD) = 2.9° (1.1°) and r = 0.99) which may be due to differences participants and methods (outlined below).

| Kinematic prediction model (Tan et al., 2022) | (Renani et al., 2021) |
|---|---|
| Participants | |
| • 17 with knee osteoarthritis | • 13 with knee osteoarthritis, 17 with total knee replacement |
| Activity | |
| • Swing and stance phases of walking, negotiating stairs, and transitioning to and from a chair | • Only walking |
| Number and location of IMUs | |
| • Two to four IMUs (unilateral or bilateral thigh and shank) | • Four IMUs (pelvis, thigh, shank, foot) |
| Validation approach | |
| • Leave-one-subject-out cross-validation | • 5-fold cross validation |

## Summary of Findings

### Kinetic Prediction – Chapter 6

**Research questions**

1. What is the prediction error of a machine learning kinetic prediction model trained on IMU data from people with knee osteoarthritis?
2. How does the performance compare between generalisable vs individualised model?
3. How does the performance compare between double-leg vs single-leg models?

**What this research adds**

This study is the first to report on the development of a BiLSTM model using IMU data for the prediction of multiple knee joint moments (flexion and adduction) and forces (compression and medial) for the stance phase of walking for people with knee osteoarthritis.

For knee joint moments the range of prediction error was nRMSE (SD) = 22% (5%) to 43% (15%), r = 0.71 to 0.8, and for forces nRMSE = 16% (4%) to 57% (23%), r = 0.78 to 0.86 for forces. For knee adduction moment, prediction error was RMSE = 0.13 (0.03) Nm/kg, nRMSE = 22% (5%), r = 0.81.

**Secondary findings**

- Individualising models by adding participant-specific training data to the generalisable model improves nRMSE by 9% to 36%.
- Double-leg models performed nRMSE 1% to 23% better than single-leg models.

**Comparison to previous literature**

Two previous studies that recruited *participants with knee osteoarthritis* had reported development of machine learning models for prediction of knee adduction moment using IMUs. He et al. (2019) reported RMSE = 0.36 Nm/kg*m and r = 0.91 using a feed-forward neural network and Wang et al. (2020) reported RMSE <0.004 Nm/kg*cm, $R^2$ = 0.947 for a decision-tree approach, and MAE = 0.004 Nm/kg*cm or <20% of average knee adduction moment, with an $R^2$ = 0.956 for an ANN approach. Direct comparison of models described in previous studies and our model is constrained by differences in normalisation procedures and reporting metrics.

Two other studies have reported on the development of ANN kinetic prediction models in *healthy participants* using identical normalisation procedures and reporting metrics to the study in Chapter 6 (Stetter et al., 2020; Stetter et al., 2019). For knee joint moments the range of prediction error was nRMSE (SD) = 18.4% to 22.3%, r = 0.71 to 0.72, and for forces nRMSE = 14.2% to 27.7%, r = 0.6 to 0.87. The strongest performing model (participant-specific, single-leg) produced lower RMSE, higher nRMSE, and stronger correlations than the two artificial neural network models (Stetter et al., 2020; Stetter et al., 2019). Considering the similarities between studies (below) the differences may be due to differences in the population or type of machine learning architecture.

## Summary of Findings

**Kinetic Prediction – Chapter 6 cont…**

**Comparison to previous literature** (cont…)

| Kinetic prediction model (Tan et al., 2022 submitted manuscript) | (He et al., 2019; Wang et al., 2020) | (Stetter et al., 2020; Stetter et al., 2019) |
|---|---|---|
| Participants | | |
| • 16 with knee osteoarthritis | • Six (He et al., 2019) and 106 (Wang et al., 2020) with knee osteoarthritis | • 13 healthy and young |
| Kinetic parameters | | |
| • Flexion and adduction moment<br>• Compression and medial force | • Only adduction moment | • Flexion and adduction moment (Stetter et al., 2020)<br>• Compression and medial force (Stetter et al., 2019) |
| Machine learning architecture | | |
| • BiLSTM | • Feed-forward neural network (He et al., 2019)<br>• Decision tree and ANN (Wang et al., 2020) | • Feed-forward neural network |
| Number and location of IMUs | | |
| • Two to four IMUs (unilateral or bilateral thigh and shank) | • One ankle IMU and six plantar pressure sensors (He et al., 2019)<br>• Two ankle IMUs (Wang et al., 2020) | • Two IMUs (thigh and shank) |
| Validation approach | | |
| • Leave-one-*subject*-out cross-validation | • Unspecified (He et al., 2019)<br>• Leave-one-out cross-validation (Wang et al., 2020) | • Leave-one-*subject*-out cross-validation |

## 7.1 Diversity Exists within the Population with Knee Osteoarthritis

Knee osteoarthritis is a condition that affects not just the knee joint, but the whole person (Caneiro, O'Sullivan, et al., 2020; Hunter, 2018; Kittelson et al., 2014). During clinical encounters, a clinician contends with substantial differences in characteristics between individuals presenting with knee pain reflecting the diversity across the population with knee osteoarthritis. Activity limitation and pain are influenced by a complex interaction of changes in physiological structure, movement patterns, psychological distress, cognitions, and alterations in neurophysiology (Kittelson et al., 2014). There seems to be consistent evidence in the literature of differences in knee kinetic (e.g. knee adduction and flexion moment), kinematic (e.g. reduced knee flexion angles), and altered muscular activity parameters (e.g. hamstring/quadriceps co-contraction) in people with knee osteoarthritis compared to people who do not have the condition (Heiden et al., 2009; Mills et al., 2013; Sonoo et al., 2019). A role of the clinician is to determine the clinical relevance of an individual's characteristics that are potentially responsive to intervention. Part of the diversity across the population with knee osteoarthritis relates to physical function, including the activities people find difficult and their affected movement patterns.

Despite diversity in presentation across the population, clinical guidelines are clear that exercise is a core intervention that should be offered to all people with knee osteoarthritis to improve clinical outcomes (activity limitation and pain) (Bannuru et al., 2019). Previous research has suggested specific baseline movement parameters are related to structural progression and clinical outcomes (Chehab et al., 2014; Hall et al., 2017; Henriksen et al., 2012; Marriott et al., 2019; Nie et al., 2019). Prior to the systematic review presented in Chapter 3, it was unclear if a relationship existed between changes in specific movement parameters and changes in clinical outcomes after exercise interventions, and therefore a review of this relationship was timely.

The systematic review presented in Chapter 3 found that across cohorts of people with knee osteoarthritis, there was only a 24.5% co-occurrence of change between knee movement parameters and clinical outcomes after exercise interventions. It is not surprising our systematic review had a similar finding to other

reviews investigating the relationship. Our results are consistent with systematic reviews of movement- or exercise-based interventions in people with knee osteoarthritis (Ferreira et al., 2015; Richards et al., 2017) as well as other body regions including the low back (Laird et al., 2012; Wernli, Tan, et al., 2020) and shoulder (Nodehi Moghadam et al., 2020). Those reviews suggest a tenuous or absent relationship between changes in movement parameters and change clinical outcomes after movement or exercise-based interventions. There are multiple reasons why studies may not have identified a relationship between changes in movement parameters that co-occur with improvements in clinical outcomes. Firstly, it must be acknowledged that there may be no relationship between changes in movement parameters and clinical outcomes after exercise interventions. But there are several possible reasons why studies may not have demonstrated a relationship if it exists. Those reasons relate to study design (e.g. statistical approaches) while others relate to heterogeneity of biopsychosocial factors within the population of those who have knee osteoarthritis (Dell'Isola et al., 2016; Roman-Blas et al., 2020). To better understand this, the findings of this systematic review will be further unpacked in the context of group-level statistical analysis, the selection of outcome measures, individual patient characteristics, and type of intervention.

### 7.1.1 Group-level Analyses

To establish a relationship between changes in two outcomes, both randomised controlled trials and cohort studies are appropriate sources of information because this research question is not about treatment effect, it is about co-occurrence. Of the studies included in the systematic review presented in Chapter 3, 20 of the studies provided only group-level data about change for both movement and clinical patient-reported outcomes, leaving just two studies that used correlation analyses to summarise this based on the strength of that relationship at an individual person-level.

Using group-level change is a blunt approach to explore the change relationship between two outcomes. Group-level change (average change) does not account for change relationships between outcomes at an individual person-level. Specifically, it does not provide information about whether the people who changed in one outcome

were the same people who changed in the other outcome. Therefore, it is unclear if those who improved in clinical outcomes are the same people who changed their movement patterns. In the systematic review when a group-level co-occurrence of change was not identified, it is possible that change relationships may have been identified had individual-level data been available and analysed using correlation.

There were only two studies that reported using individual person-level data to investigate the relationship. In comparison, a similar systematic review in people with low back pain identified 27 studies that used correlation analysis between changes in movement patterns (lumbopelvic kinematics and muscle activity) and clinical outcomes (Wernli, Tan, et al., 2020). They reported that 31% of time when a correlation was tested there was a change relationship, although most of the included studies were of low quality (Wernli, Tan, et al., 2020). Altered movement patterns are one feature found consistently in the literature in both people with knee osteoarthritis and those who have low back pain (Laird et al., 2019), therefore, similarities in individual change relationships between movement patterns and clinical outcomes may exist between these populations. However, until additional cohort studies and randomised controlled trials are available that investigate change relationships between movement patterns and clinical outcome in people with knee osteoarthritis using correlation analysis (or other individual-level approaches, such as regression), the exact nature of the relationship remains unclear.

### 7.1.2 Selection of Outcome Measures

#### 7.1.2.1 Biomechanical Assessment of Functional Activities

No studies in the systematic review investigated movement parameters for activities other than walking, therefore, only co-occurrence of change in walking-related movement parameters with clinical outcomes could be assessed. This is problematic because there is diversity in activity limitation amongst the population with knee osteoarthritis (Dobson et al., 2013; Fukutani et al., 2016; Machado et al., 2008; Roos & Lohmander, 2003). For example, in a prospective longitudinal observational study of 491 people at high risk for developing knee osteoarthritis, negotiating stairs was the activity first associated with knee pain (Hensor et al., 2015). In a study of 184 people with knee osteoarthritis over the age of 55 years, the

prevalence of having any difficulty with standing from a chair was 73% and climbing a flight of stairs was 79% (Machado et al., 2008). Those findings are supported by an earlier study of 40 participants who had knee osteoarthritis that reported an activity limitation prevalence of 50% for negotiating stairs, 16% for gait and 9% performing sit to stand (Fisher et al., 1993).

As walking was the only activity investigated, the results do not address the question about the relationship between a change in movement patterns and clinical outcomes for other clinically important activities. The search in the literature review presented in Chapter 3 discovered only one study that had investigated movement patterns and clinical outcomes during functional activities other than walking, although that study was excluded because they reported median change rather than mean change for clinical outcomes on the KOOS (McQuade & de Oliveira, 2011). After eight-weeks of machine-based strengthening exercise, they reported no group-level changes in knee joint moments and muscle activity parameters during a step-up activity despite improvement in activity limitation and pain. While that study provides some preliminary evidence suggesting an absence of a relationship during that activity, further research is required to establish if that lack of a relationship is consistent across cohorts. Further research is also warranted to establish if there is a relationship between changing movement patterns for a broader range of activities that have a high prevalence of activity limitation in people with knee osteoarthritis.

While it is important that movement parameters for other activities be investigated more broadly, there are also other considerations that may have affected our ability to identify change relationships within the systematic review. These include selection of patient-reported outcome measures, patient-specific activity limitations, if interventions are targeted appropriately, and individual participant characteristics (e.g. cognitions or baseline movement patterns).

#### 7.1.2.2 Patient-reported Outcome Measures

In the systematic review, activity limitation and pain were most frequently assessed using patient-reported outcome measures, such as the WOMAC physical function and pain subscales (Bellamy et al., 1988), or the VAS/NRS for pain (Alghadir et al., 2018). The WOMAC and its subscales are recommended as part of

the OARSI core set of patient-reported outcomes for clinical trials (McAlindon et al., 2015). When exploring the relationship between changes in movement parameters during functional activities and patient-reported outcome measures of activity limitation and pain, it is important to understand the psychometric properties of the patient-reported outcome subscales such as the WOMAC or VAS/NRS. Those properties have the potential to reduce the capacity to establish whether relationships exist between changes in movement parameters during specific activities and other clinical outcomes, as explained below. For example, the physical function sub-scale of the WOMAC is a 16-item outcome measure that provides information about patients' perception of difficulty or pain performing a range of clinically relevant physical activities rated from 0 (none) to 4 (extreme) (**Figure 7-1**). While the WOMAC is useful for assessing outcomes across the population with knee osteoarthritis, not all items are relevant to every individual.

**Figure 7-1. Items in the WOMAC subscales.**

| Activity Limitation (Physical Function) | | Pain | |
|---|---|---|---|
| **Descending stairs** | | **Walking on a flat surface** | |
| **Ascending stairs** | | **Going up and down stairs** | |
| **Rising from sitting (sitting transition)** | | At night while in bed, pain disturbs your sleep | |
| Standing | | Sitting or lying | |
| Bending to the floor | | Standing upright | |
| **Walking on flat surfaces** | | | |
| Getting in and out of a car, or on or off a bus | | | |
| Going shopping | | | |
| Putting on your socks or stockings | | | |
| Rising from the bed | | | |
| Taking off your socks or stockings | | | |
| Lying in bed | | | |
| Getting in or out of the bath | | | |
| Sitting | | | |
| **Getting on or off the toilet (sitting transition)** | | | |
| Performance heavy domestic duties | | | |
| Performing light domestic duties | | | |
| Total possible score | = 64 | Total possible score | = 20 |
| Maximum possible score for: | | Maximum possible score for: | |
| Walking | = 4 | Walking | = 4 |
| Stairs | = 8 | Stairs | = 4 |
| Sitting transition | = 8 | Sitting transition | = NA |

Of the 16 items, there is one question about walking, two about using stairs (up/down stairs), and two about moving from sit-to-stand (rising from sitting, getting on or off the toilet) (**Figure 7-1** bold). For people who only have limitations for more difficult activities, such as ascending stairs or heavy domestic duties, they may score no (0/4) difficulty or pain with walking. So, while the WOMAC is condition-specific for the population with knee osteoarthritis, the scores for each subscale can be considered as generalised, rather than activity-specific, measures of activity limitation and pain. Because there is diversity in activity limitation across the population with knee osteoarthritis, there is a potential washout effect when attempting to establish whether relationships exist between changes in movement parameters during specific activities and clinical outcomes.

Another outcome used frequently across studies in the systematic review was the VAS 0 to 10 pain scale. Like the WOMAC, the VAS asks broad questions about pain, such as rating average pain in the past week. A more specific method of assessing pain, such as pain while performing specific activities (e.g. pain intensity when ascending stairs), may help to more precisely investigate the relationship between changes in movement patterns and improvement in pain. Therefore, the selection of patient-reported outcome measures and their psychometrics could impact the likelihood that a relationship is identified. The following section discusses how a relationship between changes in movement parameters and clinical outcomes could be investigated more precisely using more personalised patient-reported outcomes.

#### 7.1.2.3 Patient-Centred Assessment

To help clinicians accommodate heterogeneous patient factors, clinical guidelines recommend individualised or patient-centred assessment for people with musculoskeletal condition such as knee osteoarthritis (Lin et al., 2020). Although those recommendations are for clinical practice, research has yet to accommodate individualised assessment of movement parameters and clinical outcomes for people with knee osteoarthritis. It is clear from the systematic review that there was an infrequent group-level relationship between walking-related movement patterns and generalised measures of activity limitation and pain. Yet, because walking was the only activity that was biomechanically analysed, and generalised patient-reported outcomes

were used, the review was not able to capture information about the relationship between change in movement patterns and clinical outcomes based on individualised assessment of activities that were clinically relevant for each participant.

For randomised controlled trials, it generally makes sense to use validated patient-reported outcome measures because they capture a broad range of important, condition-specific activity domains, across a particular clinical population. Similarly for movement parameters, studies have consistently selected only one type of activity (e.g. walking), because the activity is broadly important across a population. However, a more nuanced and patient-centred approach would be to investigate the relationship between changes in clinical outcomes and movement patterns or parameters by selecting an individual's activity-specific patient-reported items and conduct a movement analysis for that same activity. No studies have yet done this in people with knee osteoarthritis.

There have been attempts made to adapt the WOMAC physical function subscale to be more patient specific. For example, one method known as the WOMAC-top 5 is to have a patient select the top five items most important to them (Seror et al., 2008). In comparison, the Patient-Specific Functional Scale is entirely individualised, does not use predefined items, and it is sensitive to change in people with knee osteoarthritis (Horn et al., 2012). Using the Patient-Specific Functional Scale or WOMAC-top 5 instead of broader patient-reported outcome measures, like the WOMAC or VAS, may have significant implications for the ability to identify relationships more accurately between movement parameters and clinical outcomes. This would be particularly important if the intervention is also targeted toward a specific activity (which is discussed further in section 7.1.4).

The Patient-Specific Functional Scale has been used in people with low back pain to explore the relationship between movement patterns and clinical outcomes using a single-case experimental design (Wernli, O'Sullivan, et al., 2020). This type of study design collects repeated measures of an outcome prior to, during and after an intervention, which allows for cross-correlation analyses between changes in movement patterns and clinical outcomes for individual participants (Borckardt et al., 2008). That study in people with low back pain demonstrated a relationship between a

change in movement or posture and change in activity limitation or pain occurred 74% of the time for the biomechanical assessment of activities that were based on an activity limitation reported in the Patient-Specific Functional Scale (Wernli, O'Sullivan, et al., 2020). This is in contrast to the same authors' systematic review that reported a relationship 31% of the time when both movement and clinical outcomes were not individualised (Wernli, Tan, et al., 2020). More useful insight about the relationship may be found in people with knee osteoarthritis if future studies use a patient-reported outcome measure, like the Patient-Specific Functional Scale and conduct a biomechanical assessment of the same activity.

### 7.1.3 Individual Patient Characteristics

In previous sections, heterogeneity of individual activity limitation and movement patterns across the population with knee osteoarthritis were explored as potential reasons why the systematic review did not identify more frequent relationships between changes in movement parameters and clinical outcomes. That infrequent relationship may also be influenced by diversity across other patient characteristics. For example, movement patterns are affected by a complex interaction between the activity being performed, the environment, and individual biopsychosocial factors such as physiological structure, symptoms, behaviour, and cognitions (Dingenen et al., 2018; Hodges & Smeets, 2015).

Examples of behaviour and cognitions related to knee biomechanics include knee confidence and fear of movement. Worse knee confidence has been related to: increased peak knee adduction angle during the stance phase of walking (Chang et al., 2011); greater range of varus/valgus movement during mid stance (Skou et al., 2014) and higher levels of trunk flexion during walking (Hart, Collins, Ackland, Cowan, et al., 2015). Fear of movement has also been related to reduced hip abduction when side lunging (van der Straaten et al., 2020) and higher levels of trunk flexion during walking (Hart, Collins, Ackland, Cowan, et al., 2015). Therefore, consideration should also be given to how behaviours and cognitions may influence the relationship between movement patterns and clinical outcomes.

There is also diversity in movement patterns across the population of people with knee osteoarthritis and therefore assessments and interventions that are not

individualised may have limited effect on movement patterns and clinical outcomes. For example, instructions for patients performing neuromuscular exercise include moving their knee over their middle toe, which is thought to modify knee loading towards that of people without knee osteoarthritis, promoting 'optimal' alignment (Ageberg & Roos, 2015; Bennell et al., 2014). But some participants may have no, or limited potential for change due to fixed deformity of the knee. In those cases, no change in movement pattern may be possible and therefore no relationship with clinical outcomes could occur. Other participants may present at baseline with 'normal' movement patterns. Therefore, promoting 'optimal' alignment for a person with knee osteoarthritis who has normal movement patterns would likely result in no change in movement pattern, nor provide the possibility for a relationship with clinical outcomes to be present. Hypothetically, there is also the possibility that movement patterns that are considered 'suboptimal' are actually adaptive and pain relieving, or even clinically unrelated to a patient's activity limitation or pain. For example, for people with severe structural knee osteoarthritis, cross-sectional studies have reported no association (Nie et al., 2019) or negative relationships (Hall et al., 2017) between activity limitation and pain with knee adduction moment (**Table 2-2**). An individual's characteristics or combination of those characteristics may influence the relationship between change in movement patterns and clinical outcomes in response to exercise interventions.

### 7.1.4 Interventions

The vast majority (90%) of studies in our systematic review reported within-group improvements in activity limitation and pain, following an exercise intervention for people with knee osteoarthritis (see Chapter 3, Error! Reference source not found.). Changes in movement patterns were not consistently observed following generalised interventions, such as strength training and progressive walking exercise programmes (see Chapter 3, **Appendix 3-6**). Even neuromuscular exercise, an intervention that is specifically designed to reduce knee adduction moment, did not reduce knee adduction moment (see Chapter 3, **Appendix 3-6**).

One possible reason that neuromuscular exercise and other general exercise interventions (e.g. walking, strength training) did not appear to consistently change

walking-related movement patterns was because the interventions did not specifically target changes in movement parameters during walking. In contrast, targeted interventions like gait retraining, that directly targeted a change in movement pattern during the activity of walking, was associated with consistent and predictable changes in movement patterns within the systematic review (see Chapter 3, **Appendix 3-6**). For example, first peak knee adduction moment changed in the expected direction after toe-in gait modifications (Richards et al., 2018; Shull, Shultz, et al., 2013; Shull, Silder, et al., 2013; Simic et al., 2013), whereas second peak knee adduction moment changed after toe-out modifications (Hunt et al., 2018; Hunt & Takacs, 2014; Simic et al., 2013). These predictable changes in gait related movement parameters may have resulted because gait retraining directly involves a modification of a movement pattern for the same activity that is undergoing biomechanical analysis.

The concept of prescribing specific exercise that closely resembles the activity being biomechanically analysed has also been investigated in small pilot study of 13 participants with knee osteoarthritis (Thorstensson et al., 2007). They investigated the effect of eight-weeks of neuromuscular exercise on knee adduction moment during walking and single-leg rise from a chair. Knee adduction moment reduced for a single-leg rise, but not for the stance phase of walking (Thorstensson et al., 2007). These results may have occurred because a single-leg rise is functionally more similar to exercises included within neuromuscular exercise programmes (e.g. a split squat) compared to walking. Future studies may clarify if neuromuscular exercise is more effective to change movement patterns for activities like sit-to-stand that more resemble a squatting movement compared to the activity of walking that does not resemble any movement described in common neuromuscular exercise programmes.

Another possibility for the infrequent relationship between changes in movement parameters and clinical outcomes is that the interventions were not individualised to each participant's personal characteristics. While some studies included in the review suggest the interventions were individualised, the aspect of the intervention that was individualised was the exercise dose, rather than their baseline movement pattern or other individual participant-specific characteristics. If individualised assessment of baseline movement patterns is found to be clinically

meaningful in a proportion of the population with knee osteoarthritis, interventions targeting movement patterns may demonstrate superior outcomes compared to non-targeted interventions. In the literature review in Chapter 3, only one small randomised controlled trial (n = 20), investigated the effect of gait retraining individualised to each participant's baseline movement patterns (Cheung et al., 2018). Participants were instructed to freely change their movement pattern by altering their foot progression angle, hip, and trunk position, while walking on an instrumented treadmill that provided real time visual feedback changes in knee adduction moment curve compared to baseline. No other study included in the review screened participants for their baseline movement patterns to guide the most appropriate exercise approach. Interestingly, this study by Cheung et al. (2018) demonstrated the largest within-group effect sizes across all forms of exercise for changes in first peak knee adduction moment, activity limitation and pain. Another baseline characteristic that may influence how a clinician may individualise exercise prescription is radiographic alignment. For example, another study included in the systematic review (Lim et al., 2008) reported greater improvement in pain for people who had a more neutral alignment at baseline compared to those who had varus deformity after a 12-week quadriceps strengthening programme. Together these studies may suggest that individualised and targeted interventions might optimise the outcome when underlying participant characteristics and outcomes are considered. However, that has yet to be tested thoroughly in people with knee osteoarthritis.

While the review search criteria were limited to include only exercise interventions, individualised patient-centred care also may provide a greater potential to change movement patterns and improve clinical outcomes, a concept that is explored in later sections. For example, addressing biopsychosocial factors like high BMI, fear or confidence have the potential to influence movement patterns and clinical outcomes (O'Sullivan et al., 2018; Preece et al., 2021).

#### 7.1.4.1 Kinesiopathological and Symptom Modification Approaches

There are two dominant theoretical approaches to changing movement patterns. One theory is a kinesiopathological approach (Lehman, 2018) that aims to normalise altered movement parameters that are related to structural progression (e.g. knee

adduction moment). The other theory is a symptom modification approach where movement patterns are targeted for their potential capacity to change symptom response when performing an activity regardless of biomechanical risk factors for structural progression and without focus toward predefined 'normal' movement parameters (Lehman, 2018). The two approaches are not diametrically opposed. They have consistent aims about improving patients' clinical outcomes over the course of treatment, however, the direction of change in movement parameters may or may not align.

Interventions such as gait retraining (Hunt et al., 2018) and neuromuscular exercise (Ageberg & Roos, 2015) are examples of kinesiopathological approaches. While those approaches have demonstrated efficacy to improve activity limitation and pain (Bennell et al., 2010; Bennell et al., 2014; Hunt et al., 2018), it has not been demonstrated that that modifying movement patterns slows or prevents further structural changes. In the studies that target movement patterns based on a kinesiopathological approach, it is theoretically possible for some participants, that the modified movement patterns aggravated symptoms and/or worsened activity limitation, which would potentially explain the absence (on average) of a relationship.

Alternatively, the symptom modification approach targets a change in movement patterns that is dependent on a reduction in symptom response (Lehman, 2018; O'Sullivan et al., 2018; Preece et al., 2021). While the aims of many of the studies included in our systematic review were to reduce knee adduction moment and reduce pain, no study reported methods that monitored if changes in movement patterns established symptom control when performing painful activities. Targeting movement patterns for the purposes of symptom control has not been investigated in large-scale randomised controlled trials. Targeting symptom control through a change in movement pattern has been tested in a small pilot study of 11 people with knee osteoarthritis (Preece et al., 2021), which shows promising results. That study, however, does not provide evidence that changing movement patterns is necessary or sufficient to modify symptoms as there was no objective assessment of any movement parameters. Earlier in this thesis, a single-case experimental design study in people with low back pain was discussed, highlighting the high frequency of a

relationship (74%) across participants (Wernli, O'Sullivan, et al., 2020). That higher frequency of a relationship could be because the assessment of movement parameters and measures of activity limitation (Patient-Specific Functional Scale) were collected at an individual person-level, or that the activity that was biomechanically analysed aligned with the specific activity limitation (Wernli, O'Sullivan, et al., 2020). An additional factor may also be the intervention. One part of the intervention (Cognitive Functional Therapy) aims to achieve symptom control through targeting a change in individualised movement patterns for activities related to activity limitation (O'Sullivan et al., 2018). Therefore, alignment of the outcome measures and intervention may have also contributed to stronger relationships. Stronger relationships between changes in movement parameters and clinical outcomes may be found if an intervention is focused on controlling symptoms through targeting patient-specific activity limitation via modifying movement patterns for activities that align with their specific activity limitation. Biopsychosocial interventions like Cognitive Functional Therapy also target other aspects of an individual's presentation like confidence and fear avoidance which can influence movement patterns (Chang et al., 2011; Hart, Collins, Ackland, & Crossley, 2015). Integrated biopsychosocial approaches may therefore have greater potential to impact movement patterns and clinical outcomes than unimodal interventions like exercise. Further investigation is required that more closely considers individual patient biopsychosocial characteristics to clarify the clinical significance of changing of movement patterns in people with knee osteoarthritis.

There is a diverse array of patient-specific characteristics that influence movement patterns and clinical outcomes in people with knee osteoarthritis. Clinicians and researchers may benefit from adopting technology-assisted data-driven advancements, to help clarify the relevance and importance of individual patient's movement patterns and other characteristics that could help shape and deliver highly individualised interventions. This concept will be further explored in the following sections.

### 7.1.5 Summary

The systematic review presented in Chapter 3 provided evidence of an infrequent relationship between changes in walking-related movement patterns and clinical outcomes when analysed using group-level data. Infrequent relationships between changes in movement patterns and clinical outcomes after exercise are potentially explained by the study design and participant characteristics within the included studies. It is unclear if similar findings would occur across other activities like negotiating stairs or transitioning to and from a chair as they have not yet been investigated. Group-level relationships do not account for individual-level change relationships, which would be best investigated using correlation analysis or similar. The type of patient-reported outcome measures to assess activity limitation and pain used within the included studies may have diluted the ability to establish a relationship with a change in movement patterns due to lack of individual specificity. To investigate relationships between change in movement patterns and clinical outcomes more precisely, biomechanical and patient-reported outcomes could be individualised and specific to personal activity limitation or pain associated with those activities. Unimodal interventions, such as exercise, may not be sufficient to address the complex biopsychosocial influences on an individual patient's movement patterns. Interpreting the relevance of highly individual, complex biopsychosocial influences on movement patterns, activity limitation and pain may be facilitated using technology-driven approaches.

## 7.2 From Small Data to Big Data

Individualising assessment and intervention strategies that consider the complexity of the relationship between a patient's biopsychosocial characteristics, the activity being performed, and the environment in which it is performed, may provide further insight into the relationship between movement patterns and clinical outcomes. Technology-driven advancements may help facilitate assessment of movement patterns in both clinical and free-living environments. Integration of technology assisted assessment may include IMU 'big data' and 'small data' approaches that could help shape and deliver individualised care. From small data that provides information about an individual (n = 1) (Hansen et al., 2014; Hekler et al.,

2019), to big data that provides information about a population (e.g. people with knee osteoarthritis) – artificial intelligence (including machine learning and deep learning) is poised to revolutionise how data is used to facilitate clinical decision making in healthcare (Lin, 2022).

Small data can be used for monitoring individual outcomes to inform individualised interventions, and to help create big data models. On the spectrum from small to big volume data – a patient's individual health outcomes are small data, whereas databases, electronic health records and clinical registries are examples of big data. Somewhere between sits cohort studies, randomised controlled trials and machine learning models based on relatively small population samples. While a person's health outcomes are small data pertaining only to that individual, their data has a high degree of dimensionality and is influenced by interactions between biopsychosocial factors.

Big data was originally described as consisting of at least one of three core characteristics – volume, velocity and variety (Kitchin & McArdle, 2016; Laney, 2001). Volume describes the size of the data; velocity the frequency of generation or handling, recording, and publishing data; and variety the dimensionality or structure. However, it is unclear at what point data becomes 'big' (Kitchin & McArdle, 2016). When building machine learning models, each additional participant's data makes the data 'bigger' and therefore potentially more generalisable across the population.

Kongsted et al. (2020) suggested that models of care in the future will require ubiquitous technology for data capture as well as data storage solutions and data handling capabilities such as artificial intelligence. Alone, technologies like IMUs are unable to independently provide improvement in patient management. Machine learning models that use IMU data may be trained on a sample of participants prior to model development or be based on a dynamic prediction model that continually adds new participants data (or types of data) to a database, refining the model over time (Kongsted et al., 2020). Both those approaches use big data, requiring large volumes of data across a sample of the population to build a model. Big data approaches may be useful to build generalisable models for the population with knee osteoarthritis.

The IMU data collected for machine learning models in Chapters 4 to 6 consisted of multiple individual participant's small data, albeit that each person contributed thousands of multi-dimensional data points, that were combined to build small scale population-based machine learning prediction models. Combining small datasets from multiple research participants aims to train a machine learning model to be generalisable across the population on which it was trained. Taking this a step further, where IMU data is collected, structured in a standardised manner, and uploaded to a database to refine a machine learning model, it would be possible to create an even bigger, constant, and ever-growing dataset. A database containing enough IMU data from across the heterogeneous population with knee osteoarthritis could provide two opportunities – development of generalisable models trained on population data, or adaptive models that are highly individualised to participant-specific characteristics of individual test participants (see section 7.3.4.3).

It may be possible to develop generalisable, population-based machine learning models from a sufficiently large dataset for the entire population (or at least large subgroups) with knee osteoarthritis. Because leave-one-subject-out cross-validation trains a model on all the data except for one test participant, the machine learning models in Chapters 4 to 6 represent the test accuracy of a population-based model defined by the inclusion and exclusion criteria. In addition to the IMU data from each participant, the machine learning models in Chapters 4 to 6 included additional participant-specific information about which leg was affected, or in the case of bilateral symptoms, the most affected leg. Prior to the study presented in Chapter 6, it was unclear if additional small data that included the individual's specific characteristics (i.e. the participant-specific model) would help further reduce prediction error for people with knee osteoarthritis. Individualising prediction models may be important as there are differences in movement parameters between people of differing knee osteoarthritis grade, sex, BMI, age, and psychological factors which may influence the precision of IMU prediction models (Chang et al., 2011; Favre et al., 2014; Segal et al., 2013; Verlaan et al., 2018). Those biopsychosocial or individual factors could be investigated as additional training data for a prediction model to determine if there is benefit on machine learning model performance.

Because of this significant heterogeneity across the population with knee osteoarthritis, more refined models may be required for specific subgroups or phenotypes (Dell'Isola et al., 2016; Dell'Isola & Steultjens, 2018). There may be subgroups within the population with knee osteoarthritis where the accuracy of human activity recognition and biomechanical prediction models would benefit from the addition of other data that could affect a person's movement patterns. A sufficiently 'big' dataset may help to improve human activity recognition and biomechanical prediction models in the future, whereby a personalised model can be built from population data that most closely resembles an individual's profile. It is conceivable that in the future, an automated, cloud-based, adaptive model building system could train and validate models that are suited to subgroups of the population or are individualised based on specific characteristics of the user, a concept that is explored later in this chapter. Some IMU manufacturers, such as DorsaVi (Melbourne, Australia), routinely store a cloud-based copy of all data collected using their sensors for such potential purposes.

But in absence of a large database that includes IMU and potentially other clinically relevant information, a generalisable model can be refined for an individual with the addition of training data from a test participant. The results presented in Chapter 6 demonstrate improved prediction error of the model for a knee moments and forces when a generalisable model is 'individualised'. That is, when some of the test participant's data is added to the training dataset of the generalisable model, that dataset then includes data from all participants as well as some of data from the person on which the model is also being tested. While this approach improved the prediction error substantially (nRMSE 9% to 36%, **Table 6-3**), the required level of clinically significant accuracy has yet to be established (see section 7.3.4).

How a person moves is only one part of a clinical assessment. Additional aspects of a patient's presentation might be combined with biomechanical data, providing higher levels of dimensionality (increased variety) for bigger data approaches to facilitate a clinician to provide precision medicine. Big data approaches could help identify patient subgroups or phenotypes that may be more responsive to an intervention, or combinations of interventions, to help guide clinical decision

making (Ahmed et al., 2020). To individualise care, machine learning systems using IMU data for human activity recognition and/or biomechanical prediction could assist with clinician- or self-monitoring to facilitate behaviour change interventions (Patel et al., 2015). Monitoring a patient's response to an intervention using IMU data collected in clinical or free-living environments, could help inform clinical decision making and help improve motivation and intervention adherence (see section 7.4.4) (Papi et al., 2015). In the future, it may be possible that a clinician could upload an individual patient's data into an online artificial intelligence system trained on high dimensional big data (i.e. deidentified data from population based electronic health records) which compares the patient's individual, small data, to the broader population to assist with clinical decision making. The discussion below will focus on how machine learning prediction models that use IMU data can assist with monitoring movement of an individual with knee osteoarthritis to inform clinical reasoning.

## 7.3 IMU-Machine Learning Models for People with Knee Osteoarthritis

The three studies presented in Chapters 4 to 6 provide novel and clinically relevant approaches for the use of machine learning for handling IMU data in people with knee osteoarthritis. Data provided by those models could be used to provide clinically important information to facilitate clinical reasoning. Those models were developed to overcome clinical limitations for the assessment of physical function described in section 2.3.2 and to overcome the calibration and electromagnetic interference limitations of IMU. The three IMU prediction models were designed for (a) human activity recognition, (b) kinematic prediction, and (c) kinetic prediction, using deep learning approaches that could be potentially used in an automated data handling pipeline. The machine learning prediction models presented in this thesis provide the groundwork for further development of a data handling pipeline that could eventually be implemented in clinical practice to aid the assessment of activity avoidance (movement quantity) and movement patterns (movement quality) in people with knee osteoarthritis.

A recent systematic review has identified that IMU models for people with knee osteoarthritis have been studied independently (Kobsar et al., 2020), without consideration about how they would be used in a data handling pipeline for data collected in free-living environments. A handling pipeline for IMU data intended for use in people with knee osteoarthritis has previously been investigated in only one study (Emmerzaal et al., 2020). However, that study validated their human activity recognition machine learning model on healthy participants, which may have contributed to implementation issues when their system was used in people with knee osteoarthritis. In comparison, the machine learning models in this thesis were validated on people with knee osteoarthritis. While the pipeline of machine learning models in this thesis have yet to be tested in a clinical context, the series of prediction models were developed with the consideration of what information a clinician might find useful to inform clinical reasoning. The concept of a machine learning data handling pipeline and potential clinical implementation will be further explored in later sections.

The three studies in Chapters 4 to 6 have substantially contributed to the limited number of peer-reviewed publications that have trained and validated machine learning models using data collected from people with knee osteoarthritis (He et al., 2019; Renani et al., 2021; Renani et al., 2020; Wang et al., 2020). We developed the first machine learning model for human activity recognition that was trained and tested on people with knee osteoarthritis (Chapter 4). The selected activities were based on clinical guidelines for assessment of physical function (Dobson et al., 2013) while movement parameters were based on previous research suggestive as being risk factors for structural progression (Chehab et al., 2014) or because of relationships with baseline clinical outcomes (Hall et al., 2017; Henriksen et al., 2012; Marriott et al., 2019; Nie et al., 2019).

Prior to the study presented in Chapter 5, all machine learning models based on IMU data for the prediction movement parameters in people with knee osteoarthritis have been only for the activity of walking (He et al., 2019; Renani et al., 2021; Renani et al., 2020; Wang et al., 2020). This pattern in the research is consistent with the results of the systematic review presented in Chapter 3, where walking was

the sole focus of biomechanical analysis in studies that also looked at clinical outcomes. While the field of research investigating these machine learning models for clinical populations is in its infancy, people with knee osteoarthritis have difficulty with a range of activities. Therefore, there is a clinical imperative that future studies should include models or data handling pipelines that provide data about a broader range of activities other than walking.

Renani et al. (2021) were the first to develop a sagittal plane angular kinematic machine learning model for walking, trained and tested on IMU data from people with knee osteoarthritis (n = 13) and those who had received a total knee replacement (n = 17). Their bidirectional LSTM model was designed to predict triaxial knee and hip angular kinematics (flexion/extension, abduction/adduction, and internal/external rotation) for the sole activity of walking. In contrast, our approach in Chapter 5 included one movement parameter (knee flexion angle) for multiple activities. Our approach is generalisable across activities, while the approach by Renani et al. (2021) is generalisable across movement parameters for only one activity. Renani et al. (2021) reported a mean (SD) RMSE of 2.9° (1.1°) and r = 0.99 for knee flexion/extension angle, compared to our walking results of RMSE 7.04° (2.6°) to 9.7° (3.86°), r = 0.85 to 0.98. While the performance of the model by Renani et al. (2021) initially seems stronger, several factors may have influenced the results including the type of participants (knee osteoarthritis and knee arthroplasty), differences in the number of activities included in training data, validation approaches and training sample sizes. Those factors require further consideration in future head-to-head testing between the multiple activity approach used in our study, compared to the multiple movement parameter approach proposed by Renani et al. (2021). Future studies should also extend on the initial work by Renani et al. (2020) for spatiotemporal gait parameters such as including step and stride length, step width, toe out angle, cadence, stance, and swing times.

Machine learning prediction of knee adduction moment from IMU data has previously only focused on knee adduction moment during the stance phase of walking (He et al., 2019; Wang et al., 2020). The study in Chapter 6, also describes a model for the stance phase of walking but expands on these studies by providing

predictions for knee flexion moment, compression force, and medial force. It is not yet clear if these models for moments and force during the stance phase of walking are suitable for other activities. There are significant resource limitations in motion analysis laboratories in conducting research into kinetic analysis for other activities, especially using stairs. While we were able to calibrate our Vicon camera volume for a three-step staircase, we did not use a staircase with integrated force plates, precluding our ability to have a reference standard to be able to develop kinetic prediction models for negotiating stairs. Custom designed staircases with integrated force plates have been described (Whatling et al., 2010), but most motion analysis laboratories do not have access to these systems. Therefore, without the gold-standard reference standard, machine learning kinetic prediction models for negotiating stairs use do not seem likely to be developed in the near future. This may impact clinical utility of machine learning kinematic prediction models for patients presenting with stair related activity limitation.

To ensure individualised assessment is possible, future machine learning prediction models based on IMU data should be developed across a range of kinematic and kinetic movement parameters for a range of clinically important activities. Implementation in clinical practice would also require a user interface to allow selection of specific movement parameters for selected activities allowing a high level of individualisation. Advancement towards clinical implementation requires consideration of a broad range of factors, from current clinical practice to up-to-date technical knowledge of machine learning architecture and optimisation of an end user interface. This would require co-design input from important stakeholder groups to improve user experience including patients, clinicians, clinical researchers, data scientists and end-user application developers (graphic designer/software engineer) (Marvel et al., 2018; Noorbergen et al., 2021). Two factors that may help to minimise perceived barriers in adopting IMU technology in clinical practice include the development of appropriate user interfaces and ensuring outcomes are clinically important and useful to inform clinical decision making (Papi et al., 2016). One study has investigated usability and utility of an IMU interface for people with knee osteoarthritis that was designed for the purposes of human activity recognition and

knee joint loading (Emmerzaal et al., 2020). An iterative process was used for the development of the final user interface informed by feedback from clinicians and patients. They reported good usability of the patient interface, but poorer clinician perceived utility, highlighting the need for strong interface design supported by a robust human activity recognition system validated on the intended end user population (see section 2.5.5 and 2.5.7). Additional co-design work is required that includes a range of stakeholders. Having representatives for all stakeholder groups would more likely result in a user friendly and accessible interface, with greater clinical utility, that integrates a pipeline of machine learning models in like those presented within this thesis.

### 7.3.1 Bias and Limitations of Machine Learning Studies

Artificial intelligence approaches are set to disrupt healthcare (Jiang et al., 2017) although care is required in its design as there is increasing recognition that artificial intelligence is inherently biased (Parikh et al., 2019). Machine learning prediction models are only as good as the data on which they are trained and are affected by unintended and societal biases (e.g. racism, weight stigma, sexism) (Parikh et al., 2019). Machine learning in healthcare has the potential to be affected by missing data, inappropriate sample sizes and classification or measurement error that may unintentionally adversely impact vulnerable populations (Gianfrancesco et al., 2018).

The machine learning models reported within this thesis were trained on data from a subset of the diverse population who have knee osteoarthritis. The participants appeared to have a BMI, knee osteoarthritis grade, activity limitation, and pain, that were representative of populations described in the majority of randomised controlled trials (Fransen et al., 2015) and cohort studies. Therefore, the machine learning models may be of use in future clinically oriented studies, and for patients who meet the same inclusion criteria. However, because there were inclusion and exclusion criteria, there is potentially variability in how the prediction models would perform across populations that differ from those in the sample. For example, participants were 60% male, predominantly Caucasian, with an average BMI ~25 $kg/m^2$, and age of 66 years. It has been reported that there are sex-specific (Segal et al., 2013), race-specific (Sims et al., 2009), BMI-specific (Verlaan et al., 2018), and age-

specific (Favre et al., 2014) differences in movement patterns in people with knee osteoarthritis. Therefore, further investigation needs to explore if these factors need to be included in models to optimise performance of human activity recognition and biomechanical prediction models for people with knee osteoarthritis. Consideration of these factors is of critical importance to minimise cultural biases in future machine learning models for people with knee osteoarthritis. It is also unknown how the models within this thesis perform across the different levels of self-reported activity limitation, pain intensity and radiological grade. Those factors have the potential to also impact the results and requires exploration in future studies to identify any such limitations of machine learning models that use IMU data.

Performance of machine learning prediction models may be affected by the location where the data was collected. IMU data used to train the models in Chapters 4 to 6 were collected in laboratory conditions. Therefore, there is uncertainty about how the models will perform when tested on data collected in free-living environments. Considering there is some indication in the literature that people move differently when observed compared to when they are in free-living environments (Brodie et al., 2016; Hillel et al., 2019), machine learning models trained on laboratory data may have diminished validity in free-living environments. For example, in 20 young healthy participants, Gyllensten and Bonomi (2011) reported a drop in prediction accuracy from 92% to 75% for an activity recognition neural network that used waist-mounted accelerometer data collected in a laboratory compared to a home environment. Future studies must consider training models on data collected in free-living condition or validate laboratory-based models in free-living environments.

### 7.3.2 Selecting Appropriate Activities

There is no consensus about the type and number of activities that should be included for training phase machine learning prediction models. Ideally, if a model is generalisable, the selected activities should be broad enough in scope to be relevant for the majority of the population with the condition, while sufficiently narrow that data provided to a clinician is clinically meaningful. The models presented in Chapter 4 to 6 were developed for the activities of walking, negotiating stairs and transitioning to and from a chair, as they are thought to be the activities most commonly affected

in people with knee osteoarthritis (Dobson et al., 2013). However, people perform a range of other activities (e.g. performing exercise, walking on an incline, and stepping over an object, turning around a corner) which may require modifications to the human activity recognition classifier. But including additional activities may affect model performance.

A system that is individualised would need to accommodate other activities by including only activities of interest as part of training the machine learning model and excluding activities that are not of interest at the point of classification. Use of a sufficiently large 'big data' database that includes a range of potentially clinically important activities may help avoid that limitation. Including other activities would require a significant amount of additional data to be collected or alternatively the addition of simulated or augmented data (Renani et al., 2021) (see section 7.3.4.1). To train individualised models only on selected activities of interest from a larger cloud database (see section 7.3.4.3) may help to limit the number of activities on which the model is trained, preserving model performance. Excluding activities outside of the training dataset at the point of classification would require those activities to be labelled as 'unknown' based on an algorithm that evaluates the probability that the wearer was not performing any activities within the training dataset (Emmerzaal et al., 2020). Both big data and exclusion-based methods are likely to be important for implementation but have yet to be tested.

### 7.3.3 Minimising the Burden for Patients and Clinicians

The foundation for obtaining clinically useful information from machine learning models that use IMU data has been established across knee osteoarthritis and other health conditions (He et al., 2019; Rast & Labruyère, 2020; Tan, Beheshti, et al., 2021; Tan et al., 2022; Wang et al., 2020; Wang et al., 2021). Machine learning architecture will continue to be optimised for specific clinical conditions, and ongoing academic and commercial research focused on improving hardware and software will influence the costs and access of integrating technology into clinical practice. However, adopting new technology comes with the burden of learning the benefits and limitations of a system.

For machine learning based on IMU data to be adopted into routine clinical practice, it is essential that developers of this technology minimise the burden for both patients and clinicians. Accounting for bias within and limitations of previous studies will help to reduce this burden. Other factors that may influence clinicians adopting IMU-machine learning systems include data handling and processing requirements, the number of IMUs required, and IMU placement limitations.

#### 7.3.3.1 Automated Data Segmentation and Labelling

IMU datasets can be very large, especially if collected over many hours from multiple sensors. A recent systematic review by Kobsar et al. (2020) has found that studies using IMU in people with knee osteoarthritis have not developed integrated data handling pipelines for human activity recognition and biomechanical outcomes. This is of concern as processing IMU data in clinical practice can be time consuming, particularly so for unlabelled data collected in free-living environments. The series of machine learning models presented in this thesis were designed to be used sequentially as an automated data handling pipeline to provide clinically important information. However, the proposed pipeline still requires proof-of-concept testing.

Human activity recognition and biomechanical prediction models that are designed to be integrated assist in minimising the time burden for clinicians by removing the need for them to develop expertise and spend time processing and labelling IMU data. Our CNN human activity recognition model in Chapter 4 was designed with a sliding window data sampling approach (**Figure 4-5**) which provides the opportunity to output the start and end times of an activity or phase of an activity and therefore label an activity in a long data stream. This provides an automated method of segmenting the data for subsequent biomechanical analysis and is an alternative method to gait event detection (Fadillioglu et al., 2020). However, we did not use the human activity recognition model for segmentation of data in subsequent studies. A true test of clinical utility for an integrated system will require exploration of whether a data handling pipeline can be created using the human activity recognition model to effectively segment data for subsequent biomechanical analysis. While this is theoretically straightforward and intuitively the next step, it is untested.

#### 7.3.3.2 Technology Needs to Be Convenient

Convenience is a key determining factor related to adoption of technology by health care workers (Arkorful et al., 2020), and a key practical issue that may affect compliance wearing sensor based technology. Prediction models requiring fewer IMUs would theoretically improve the patient and clinician burden as fewer IMUs would be more discrete, reduce the time burden of organising adhesive and placing IMUs accurately, as well as reduce data processing time.

We elected to place IMUs on the thigh and shank to explore the performance of single-leg compared to double-leg models. In comparison, Renani et al. (2021) trained their kinematic prediction model on four IMUs placed on the pelvis, thigh, shank, and foot for walking. They demonstrated lower prediction error than both the single- or double-leg models, possibly suggesting that four IMUs may be better placed across different body regions rather than across two legs. However as other explanations may account for those differences (e.g. classifying multiple activities vs only walking), further investigation is required to explore the best number and location of IMUs in head-to-head trials.

While fewer IMUs is more convenient, a potential trade-off exists between the number of IMUs and a model's prediction error. In Chapters 5 and 6, we explored the effect on prediction accuracy of two different types of models that included input data from different numbers of IMUs. The single-leg model was trained on data from two IMUs placed on the leg of interest, while the double-leg model included training data from two additional IMUs on the contralateral leg. In Chapters 5 the single-leg sagittal plane angular kinematic model demonstrated lower prediction error than the double-leg model for the activities of walking or negotiating stairs. We concluded that a single-leg model (2 IMUs on one leg) may be more appropriate for ambulatory activities where each leg is moving asynchronously – for instance, when one leg is swinging the other is in stance. In direct contrast, for sit-to-stand and stand-to-sit the double-leg model was more accurate. However, the differences in prediction error between the single-leg and double-leg kinematic prediction models were small and potentially not clinically significant. Therefore, two IMUs on a single leg may be preferable because of the convenience factor, compared to placing twice as many

IMUs and processing twice as much data. As IMU based prediction models evolve, such design trade-offs will be an on-going consideration.

While we found that single-leg models had less prediction error for walking in our kinematic model, it was unclear prior to the study in Chapter 6 if double- or single-leg models were more accurate for kinetic prediction. The results in Chapter 6 suggest that double-leg (LOSOCV) models are up to 23% more accurate for medial force predictions, 6% more accurate for knee adduction moment, but only 1% to 2% different for compression force and flexion moment (**Table 6-4**). This may suggest there is a stronger relationship between how each leg moves in the coronal plane than for other planes of movement. Also, when individualising the model, there was only 1% to 2% difference in prediction error for coronal plane movement parameters, which may suggest that across the population with knee osteoarthritis, there may be higher variability in the coronal plane movement parameters. This has two significant implications for future development of kinetic prediction models for medial force and knee adduction moment, (a) for generalisable models double-leg models may be preferred, and (b) if individualised models are developed, single-leg models may be sufficient. Further studies are required to compare model performance using those approaches and explore convenience for the user using a greater number of IMUs compared to collecting data within a consultation to help train an individualised model.

The most convenient situation would be to have a single IMU that provides sufficient information to train a machine learning model with minimal prediction error. Machine learning models have been trained on data from as few as one IMU for the purposes of human activity recognition and predicting movement parameters of the lower limbs. A number of studies have investigated the performance of biomechanical prediction models using a single IMU placed on different body regions on healthy participants (Coskun et al., 2015; Jiang et al., 2020; Lee & Lee, 2022; Lim et al., 2020; Sung et al., 2022). While those studies provide preliminary evidence that single IMUs can be used to predict a range of movement parameters in healthy people, it is unclear how those models perform for people with knee osteoarthritis. Future studies need to determine the number and best location of IMUs required to

maximise model performance while minimising inconvenience for different populations.

In both multiple and single IMU arrangements, IMU misplacement requires consideration. Accurate placement increases the burden and misplacement may impact the validity of the results. Future investigation needs to investigate the impact of IMU misplacement and focus on the development of machine learning prediction models using IMU data that are robust to variability in sensor placement.

Another factor that impacts user convenience is the type and number of technologies used within a prediction model. IMUs might be only one part of an integrated human activity recognition or biomechanical prediction machine learning model. For example, a range of environmental and body worn technology has been used for human activity recognition, including cameras, smart watches, smart shoes/pressure soles, global positioning systems, and Wi-Fi signals (Qiu et al., 2022). Combinations of technology may be required to optimise model performance but increasing number of technologies integrated impact user burden. To ensure convenience of use for both the wearer and clinician, the required technology for any machine learning system needs to use sensor data that is as unobtrusive as possible, has limited privacy issues, is affordable, and harnesses technology that is readily available in society.

### 7.3.4 When is Accurate, Accurate Enough?

Clinicians collect outcomes, and for those outcomes to be clinically useful, their statistical performance must be known. Clinically meaningful data can help to inform clinical reasoning. However, defining clinically meaningful performance of a machine learning model is complex. For human activity recognition there are no guidelines that determine what level of accuracy (overall accuracy), precision (positive predictive value) or recall (sensitivity) is clinically meaningful. Incorrect activity classifications will impact user burden, and therefore should be optimised. Ideally machine learning biomechanical prediction models would be as precise as gold-standard motion analysis systems. Yet error is ever present and needs to be interpreted within a clinical context. For implementation to occur, machine learning models using IMU data may need to integrate additional individual patient characteristics (e.g. BMI, sex,

psychosocial factors), include only relevant activities, potentially be individualised, have robust validation performed for data collected in free-living environments, and have methods for handling activities the system was not designed to detect. The prediction error is dependent on those factors and others including the population, intra-participant variability, inter-participant variability, the activity, and the type of movement parameter.

In biomechanical research, a common error threshold to determine clinically relevant reliability (McGinley et al., 2009) and validity (Robert-Lachaine et al., 2017; Robert-Lachaine et al., 2020) is 5°. Physiotherapists on average demonstrate a visual accuracy threshold of 12° for detecting a change in movement across body regions for single plane movement (Abbott et al., 2022). Also available are smart phone camera-based apps that when compared to the gold-standard Vicon system have been reported to have mean difference 5° (95% limits of agreement ranging -17.6° to 7.6°, and a minimum detectable change of 6°) (Krause et al., 2015). Although smart phone camera-based apps are not possible use across unrestricted free-living environments (i.e. outside the field of view).

The acceptable level of accuracy is a clinical decision rather than a statistical one, and likely to be context specific. For biomechanical predictions to have any meaning in clinical practice, the change in a movement parameter must be larger than the measurement error of the outcome measure. The degree of clinically acceptable prediction error would differ depending on the activity of interest. For example, a 5° change in knee flexion is likely to be more clinically meaningful for the stance phase during walking where the sagittal plane range of movement is relatively small compared to sit-to-stand where 5° is arguably trivial. Therefore, clinicians need to consider that the system they adopt should be at least as accurate as any expected changes in their patient's movement pattern to be confident a true change in movement has occurred.

For knee joint kinematics measured with IMUs, a systematic review by Poitras et al. (2019) reported strong evidence that the RMSE ranges between 1° to 11.5° and correlation coefficients (r) between 0.4 to 1. While the RMSE tells a clinician how

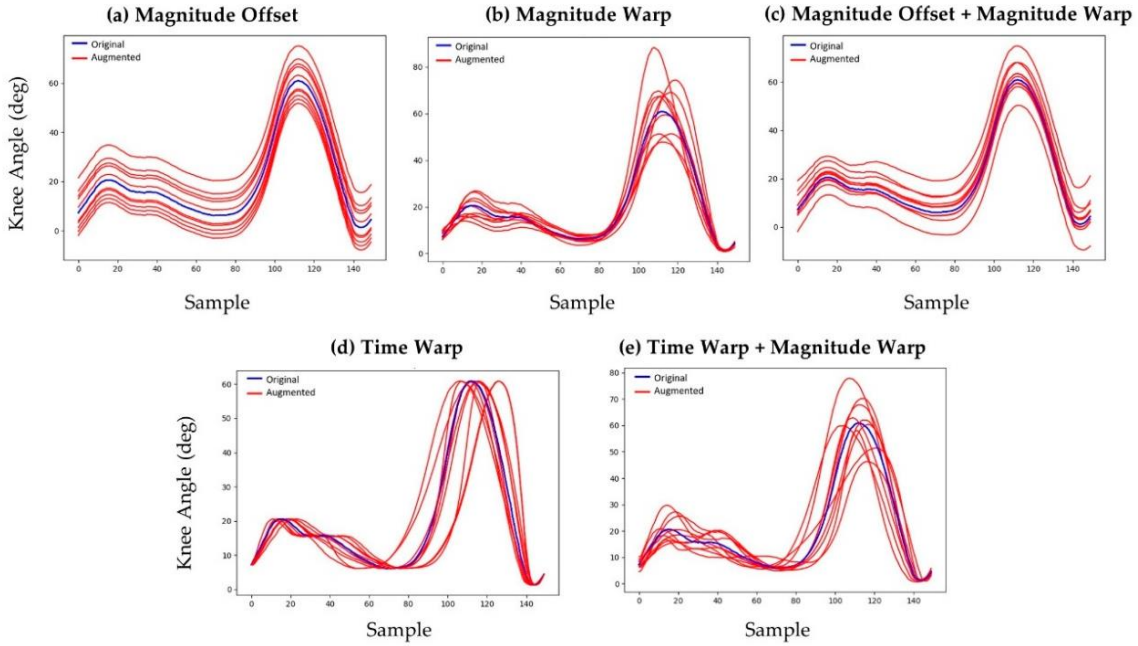
similar the prediction model is to the reference standard (the average distance between the prediction and reference standard), the r values indicate if the model has a relationship with the reference standard, which for a pattern of movement is a non-linear prediction. The pattern of movement is important if additional calculations are made in reference to the time scale, such as the rate of change of a movement parameter (e.g. degrees/second, knee adduction impulse (Nm/kg/s) – outcomes that may be of clinical interest. The results of the kinematic prediction model presented in **Table 5-3** show that the RMSE is on the higher end of the reported range of IMU studies that use traditional methods of estimating kinematics (e.g. Kalman filter) (Poitras et al., 2019). Multiple changes could be made to optimise the accuracy, including enhancing the training dataset, developing activity specific models, individualising prediction models, or refinement of machine learning architecture.

#### 7.3.4.1 Enhancing the Training Dataset

Collecting biomechanical data is time and resource heavy. We recruited 18 participants for the machine learning studies as that number was consistent with the number of participants in other studies (range n = 6 to 30) (Findlow et al., 2008; He et al., 2019; Renani et al., 2021; Stetter et al., 2020; Stetter et al., 2019; Wouda et al., 2018). With further data cleaning for studies in Chapter 5 and 6, the training dataset required two participant's data to be removed. Data cleaning is required because it is common in biomechanical studies that not all collected data is of sufficient quality due to procedural issues or technological malfunction.

The machine learning models within this thesis were trained only on the actual data collected from the IMU and Vicon (+ / - force plate) systems. Using simulated or augmented data in addition to the actual collected data is one way of increasing the sample size and variability within the training dataset. Simulating IMU data involves using the reference standard (e.g. Vicon) data to develop additional simulated samples, while augmenting data involves creating additional samples by offsetting or warping the magnitude or time domains (**Figure 7-2**).

**Figure 7-2. Examples of augmented data.**




Note. "Various data augmentation methods used to generate synthetic kinematic data: (**a**) magnitude offset, (**b**) magnitude warping, (**c**) combined magnitude offset and magnitude warping, (**d**) time warping, and (**e**) combined time warping and magnitude warping." From " The Use of Synthetic IMU Signals in the Training of Deep Learning Models Significantly Improves the Accuracy of Joint Kinematic Predictions" by M.S. Renani, A.M. Eustace, C.A. Myers, and C.W. Clary, 2021, Sensors, 21(17), p. 6. (https://doi.org/10.3390/s21175876). Copyright 2021 by the authors (Creative Commons)

Performance of IMU-machine learning models improves when trained on simulated or augmented data (Dorschky et al., 2020; Mundt, Koeppe, David, Witter, et al., 2020; Renani et al., 2021). There is some evidence that combining simulated and augmented data with real IMU training data when predicting sagittal plane angular kinematics improves prediction error by 30% (Dorschky et al., 2020) to 50% (Renani et al., 2021), but only has limited impact (6%) on kinetic parameter prediction (Dorschky et al., 2020). This more substantial improvement in kinematic prediction compared to kinetic prediction has been replicated with authors concluding that for kinetic parameters, accuracy seems to be improved with more noise from soft tissue artefact in the data than from the size of the augmented data set (Mundt, Koeppe, David, Witter, et al., 2020). Augmenting data to introduce 'jittering' to simulate noise (e.g. soft tissue artefact) in the data can improve kinetic prediction model performance (Um et al., 2017). Simulating and augmenting data helps to reduce researcher burden by reducing the number of participants required and trials collected, while augmented data has the added benefit of introducing variability into

the model which may help to improve accuracy and generalisability. Further investigation using these approaches on the models presented within this thesis may help reduce prediction error.

Another consideration is a 'data-centric' approach to building machine learning models (Ng, 2021). Traditional approaches to building machine learning models have focused on what is known as a 'model-centric' approach whereby model performance is enhanced through optimising machine learning architecture and training procedures. The data-centric approach to enhancing performance of machine learning models was proposed by globally recognised leader in artificial intelligence Andrew Ng. He suggests that while 80% of a machine learning system depends on the input data, and 20% on the machine learning model, only 1% of the research has focused on optimising data and 99% of research has focused on optimising machine learning architecture (Ng, 2021). Some aspects of a data-centric approach to developing machine learning prediction models for IMU data include focusing on consistency of data collection methods, assessing of data quality and data cleaning (Mazumder et al., 2022). While we had specific protocols and processes developed for those purposes, further investigation should focus on optimising those aspects of developing the machine learning model that would further enhance the training dataset.

#### 7.3.4.2 Developing Activity-specific Models

Most machine learning biomechanical prediction models are trained on a single activity, limiting their generalisability for use in free-living environments (He et al., 2019; Lim et al., 2020; Mundt, Koeppe, David, Witter, et al., 2020; Mundt, Thomsen, et al., 2020; Renani et al., 2021; Wang et al., 2020 ). The most straight forward method of developing a machine learning biomechanical prediction model for multiple activities would be to collect IMU data from a sufficiently large number of activities to develop a single generalisable model. Yet, there are a wide range of activities that are performed by an individual and across different populations, which may affect the validity of a model. In Chapter 5, we trained a deep learning model to predict sagittal plane angular kinematics for a range of clinically important activities for people with knee osteoarthritis (walking, negotiating stairs and transitioning to and from a chair) (Dobson et al., 2013). But it is unclear if this approach, combing all activities together

to train a model, affects model performance. The results of Chapter 5 demonstrate that when a biomechanical prediction model is trained on multiple activities the model performance is not equal across activities – a concept that is supported in the two studies by Stetter et al. (Stetter et al., 2020; Stetter et al., 2019). Participants in those studies had two IMUs placed on the thigh and shank on a single leg. For our single-leg kinematic models, we identified higher prediction error for activities where both legs are doing the same movement (e.g. sit-to-stand) compared to activities where each leg is doing different movement (e.g. walking). Consistent with our findings, previous studies reported higher prediction error for two-leg synchronous activities (e.g. jumping) compared to asynchronous activities (e.g. walking) (Stetter et al., 2019). Together these findings suggest that either double-leg models are more appropriate for symmetrical two leg activities or that separate activity specific models may be required for each activity.

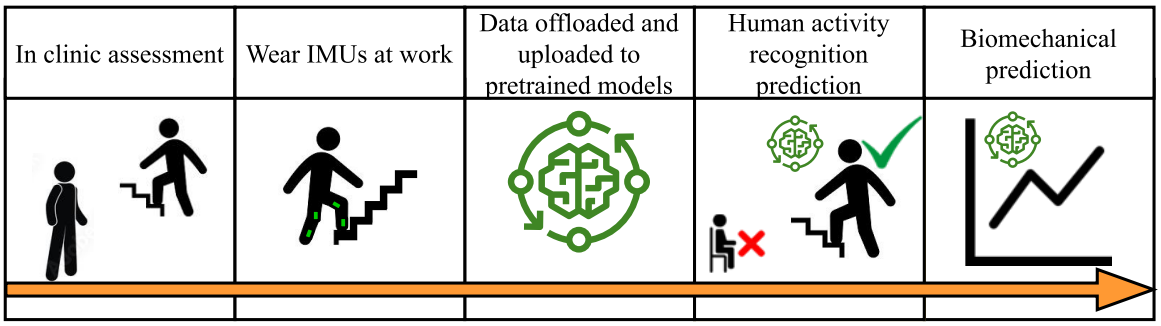
While differences in requirements of the lower limbs during different activities seem to drive the difference in model performance, this has not yet been extensively tested and requires further investigation. Further, it is possible that some movement parameters may be more accurate when trained on data that considers both lower limbs (double-leg model) rather than only the one being predicted (single-leg model), a concept that was previously explored in the previous section 7.3.3.2.

#### 7.3.4.3 Individualising IMU-Machine Learning Prediction Models

Customisable (e.g. for activities only relevant to the patient) and individualised models (i.e. enhanced by the patient's own data) may improve machine learning prediction models that use IMU data. The study presented in Chapter 6 set out to explore the effect of individualising a machine learning prediction model by adding some data from the test participant to the training of the model that included data from all other participants. We found 9% to 36% better performance across kinetic movement parameters (**Table 6-3**) when the models were individualised. This approach would presumably improve the performance of the human activity recognition and sagittal plane angular prediction models in Chapters 4 and 5, but this has yet to be tested.

Depending on how a machine learning data handling pipeline for IMU data was created, there would be different requirements on a clinician and patient to collect and process the relevant data. During a clinical encounter, one option would be to have the patient first assessed by a clinician, leave the clinic, then wear the IMUs in the required environment, after which the (small) data would be offloaded and uploaded to a local IMU system where the machine learning pipeline would provide predictions for which activities were performed and the selected movement parameters (**Figure 7-3**).

**Figure 7-3. Generalisable prediction model workflow**



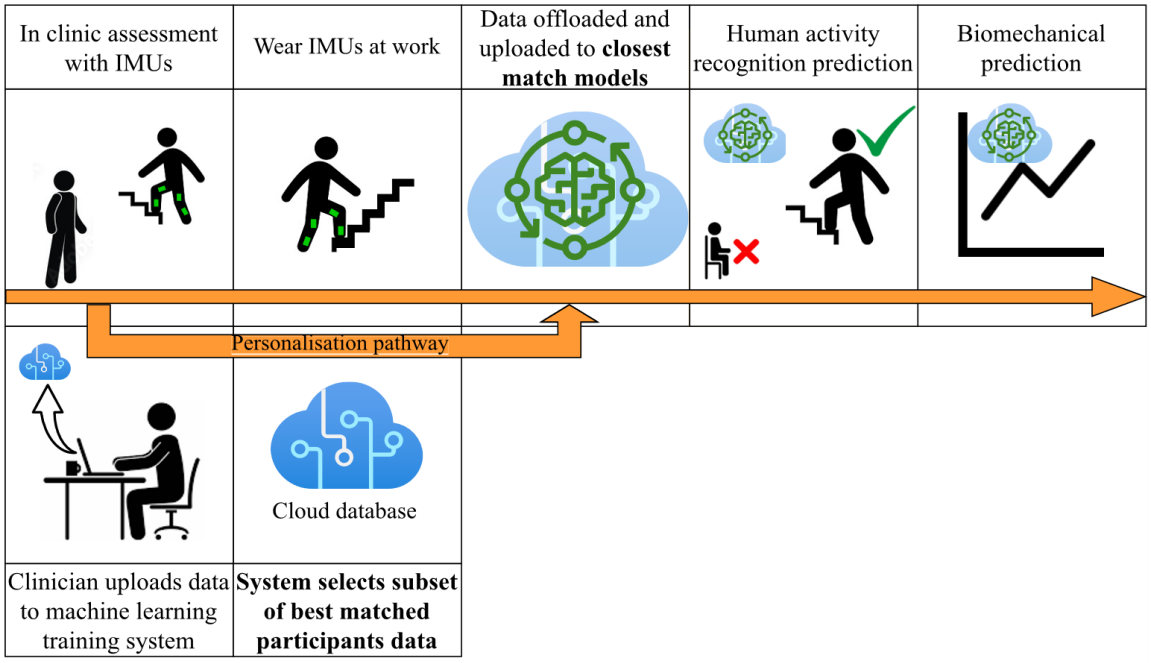

Generalisable models have previously been attractive because if trained on a sufficiently large number of the heterogeneous population, they can be applied more broadly across that population. But generalisable models have limitations for complex systems like people who have knee osteoarthritis, and the various biopsychosocial factors that affect movement patterns. Previous sections within this discussion chapter have highlighted the limitations of making clinical decisions about an individual patient from group-based data. It is clear that in a clinical setting, ideally, both clinical outcomes and measures of physical function are individualised and personally meaningful.

As the future of healthcare grows to include options for instantaneous uploading and processing of data files, one method of individualising models is to use data from a single person with knee osteoarthritis to help improve the prediction accuracy of generalisable models when applied to that same person. Personalised models have also improved prediction accuracy for other populations. For example, for people with Parkinson's disease, Rodríguez-Martín et al. (2017) developed a machine learning model to predict freezing of gait episodes, demonstrating an 11%

improvement in personalised models compared to a generalisable model using a leave-one-subject-out cross-validation approach.

However, in clinical practice there is no reference standard (e.g. Vicon and force plates), which is a significant limitation for the methods used in Chapter 6 to personalise the models. In a clinic it may be possible to use a non-gold-standard system like Microsoft Kinect for estimating kinematics (Pfister et al., 2014) as a reference standard, albeit with reduced precision than Vicon. But there seems to be no commercially available motion analysis systems that also integrate force plate data required to establish kinetic reference standards. Therefore, one solution would be to develop dynamic individualised prediction models instead of a single generalisable prediction model. A dynamic approach that aims to accommodate for heterogeneity could use a 'closest match' method to machine learning model building whereby there is a cloud database housing large datasets (big data) from a diverse range of people with knee osteoarthritis (**Figure 7-4**).

**Figure 7-4. Closest match model building based on patient's clinic data**




During the patient-clinician physical assessment, the patient would wear the IMUs, perform relevant activities and the clinician would then upload their personal small dataset to the cloud-based system that houses the database. The cloud-based system then compares the patient's data to that in the database, identifies selected

people's data from the database that is the closest match to the patient and automatically trains a new, personalised model for the patient. Theoretically, a highly personalised model would include IMU data, and possibly other patient-specific data across a range of domains may improve machine learning model performance.

If future studies demonstrate that adding highly personalised data improves performance of generalisable IMU-machine learning prediction models, big data approaches to creating highly personalised IMU-based monitoring systems may be eventually included in routine clinical practice and help inform individualised management.

#### 7.3.4.4 Health Researchers, Clinicians, and Data Scientists Working Together

The studies presented within Chapters 4 to 6 of this thesis represent an interdisciplinary approach that brought together experts in clinical practice, data science, health research, biomechanics, biostatistics, and data management. Clinicians are experts in diagnostics and the management of health conditions. They understand how a condition affects a person based on their foundational knowledge and skills, research evidence, and practical experience, which are combined with expertise in interacting with individual patients. Importantly, clinicians can provide insight into practical considerations about workflow when using technology in practice (**Figure 7-3** and **Figure 7-4**). Health researchers are experts in conducting quantitative or qualitative studies for populations with health conditions to, amongst other aims, help aid clinicians in clinical decision making. As health data continues to grow with integration of technology, data scientists bring expertise in handling large datasets to find patterns in data through various methods that can include machine learning and other statistical methods. Into the future, it is essential that experts in each of these fields work together to optimise technology driven applications like machine learning models that use IMU data.

Technology facilitated healthcare is becoming increasingly ubiquitous, but more commonly than not, health researchers and clinicians do not have the technical expertise to handle large data sets like those produced by IMUs, nor develop machine learning models. This may provide an explanation as to why most IMU-based machine learning studies cited within this thesis involved data scientists or machine

learning experts but not health researchers or clinicians. While those studies provide important foundational work that may help answer clinical questions, the focus on those studies is usually model-centric (Ng, 2021), one of optimising machine learning architecture or comparing different type of machine learning models. An alternative, more data-centric approach is to work with collaborators with diverse but complementary skill sets. Together, they can ensure that high quality, clinically meaningful training data (described earlier in this section) is collected to help design machine learning models to output information that is clinically useful. Further work is required that take advantage of knowledge from a range of stakeholders to ensure any system that is to be eventually implemented optimises not just machine learning model architecture, but also clinical and practical considerations about how the information would eventually be used.

### 7.3.5 Summary

The findings presented in Chapters 4 to 6 provide one of the earliest bodies of work describing the development of machine learning using IMU data for human activity recognition and biomechanical prediction in people with knee osteoarthritis for clinically relevant activities. Those studies build on foundational work that developed IMU prediction models in healthy populations and a limited number of studies that included people with knee osteoarthritis. Prior to this body of work, machine learning models for IMU data intended for use in people with knee osteoarthritis had been developed as discrete systems, including only the single activity of walking, and typically only for single movement parameters. The proposed pipeline of machine learning models addresses a key gap in the literature that considers data handling requirements for implementation in clinical practice, providing information about multiple clinically relevant activities and movement parameters. Prior to implementation of IMUs and machine learning data handling pipelines in clinical practice, there is a requirement for a significant amount of research to address a number of clinical and machine learning methodological questions. With further investigation, machine learning models like those presented in this thesis may provide researchers and clinicians a tool that could help assess

individualised baseline movement patterns or monitor changes in movement patterns over time to augment patient-reported outcomes and performance-based testing.

## 7.4 IMUs as part of Individualised Assessment and Management

The systematic review in Chapter 3 reported a tenuous relationship between changes in movement parameters and clinical outcomes in people with knee osteoarthritis. Section 7.1 unpacked the reasons why the studies included in the systematic review did not consistently find a relationship with one concept being the heterogeneity of patient characteristics across the population. In order to accommodate heterogeneity, there are calls for individualised assessment and management in clinical practice (Caneiro, O'Sullivan, et al., 2020; Kongsted et al., 2020; Lin et al., 2020). To facilitate individualised assessment, IMUs paired with machine learning approaches could help facilitate monitoring of patients in free-living environments. This section will explore the intersection of clinical practice, clinical biomechanical research and IMU monitoring for the purposes of individualised assessment and management.

### 7.4.1 Towards Individualised Assessment and Management

Prior to implementation of IMUs and machine learning data handling approaches in clinical practice, there is a need to establish the type of patient profiles that may benefit from monitoring of activities and movement patterns that might be amenable to modification using targeted interventions. Across the population of people with knee osteoarthritis, heterogeneity exists across many aspects of a patient's presentation (Dell'Isola et al., 2016) and response to intervention (Knoop et al., 2011). Improved treatment response may occur should heterogeneity be accounted for, although there is currently no clear evidence for this in people with knee osteoarthritis. Therefore, there have been calls for greater focus on individualised and multidimensional assessment and interventions for people with knee osteoarthritis (Caneiro, O'Sullivan, et al., 2020; Holden et al., 2021; Hunter, 2018; Hutting et al., 2022; Karsdal et al., 2014; Kittelson et al., 2014; Kongsted et al., 2020; Rausch Osthoff et al., 2018) and other musculoskeletal conditions (Caneiro, Roos, et al., 2020).

One step towards individualising assessment and management for people with knee osteoarthritis is using an approach known as phenotyping (Lane et al., 2011). Phenotyping is characterising subgroups or clusters of people based on multiple observable common data points across a population with a health condition (Dell'Isola et al., 2016; Felson, 2010; Kittelson et al., 2014; van Spil et al., 2020). There are multiple studies now demonstrating that different clinical phenotypes of knee osteoarthritis exist that have different levels of activity limitation and pain (Dell'Isola et al., 2016; Dell'Isola & Steultjens, 2018; Felson, 2010; Kittelson et al., 2014; Knoop et al., 2011; Roman-Blas et al., 2020). One of those phenotypes that is estimated to represent 12% to 22% of the population with knee osteoarthritis is related to biomechanical alterations that may respond more favourably to individualised interventions that target movement patterns (Dell'Isola et al., 2016; Roman-Blas et al., 2020).

Phenotyping in people with knee osteoarthritis demonstrates that biomechanical factors such as specific movement parameters or movement patterns may be more clinically relevant for some people, and less for others. That diversity in clinical relevance of movement patterns was not considered within the studies included within the systematic review presented in Chapter 3 in people with knee osteoarthritis, nor in a similar systematic review in people with low back pain (Wernli, Tan, et al., 2020). However, when study designs can accommodate heterogeneity, there seems to be stronger relationships between movement patterns and clinical outcomes for people with low back pain (Wernli, O'Sullivan, et al., 2020; Wernli et al., 2021). The results of the systematic review and others that demonstrated similar findings may have been different should the studies included only participants with a biomechanical phenotype.

In the future, it may be possible to assess each patient or research participant using machine learning human activity recognition and biomechanical prediction models like those presented in Chapters 4 to 6. Should the models be integrated into a data handling pipeline, a machine learning IMU system could facilitate assessment of movement patterns that match a biomechanical phenotype. Individualised assessment using machine learning-based systems of IMU data for the prediction of

knee moments and forces, like that presented in Chapter 6 could help identify baseline biomechanical risk factors for structural progression (Chehab et al., 2014). There remains the possibility that interventions that target movement patterns or knee joint malalignment could slow the structural progression in knee osteoarthritis. For example, there is evidence that people with knee osteoarthritis who have more toe out have reduced odds of structural progression over 18 months compared to those who have less toe out (Chang et al., 2007). That suggests that gait retraining (Hunt et al., 2018), an intervention demonstrated to successfully reduce knee adduction moment in Chapter 3, may be a suitable option to modify biomechanical risk factors for structural progression. Further studies should investigate if reductions in knee adduction moment is a mediator of slower structural progression in people with knee osteoarthritis. Should future research demonstrate the ability to slow progression of structural change through various movement-based or exercise interventions, machine learning systems of IMU data are well positioned to monitor improvement or worsening of biomechanical risk factors in both clinical practice and research.

Recently there are detailed descriptions about the framework for comprehensive individualised biopsychosocial interventions for people with knee osteoarthritis (Preece et al., 2021) that were adapted from Cognitive Functional Therapy for people with low back pain (O'Sullivan et al., 2018). The proposed intervention (Cognitive Muscular Therapy) described by Preece et al. (2021) is a biopsychosocial approach for the management of people with knee osteoarthritis that targets maladaptive beliefs and behavioural responses, reduction of muscular co-contraction and knee loads, and functional retraining. The approaches described by Preece et al. (2021) and O'Sullivan et al. (2018) suggest that management may target movement patterns only after sound clinical reasoning that establishes a justifiable relationship between movement patterns and activity limitation or pain. Currently, no studies have used a comprehensive individualised assessment using IMU data as the key input for individualised management and provided details about changes in both biomechanical and clinical outcomes in people with knee osteoarthritis. Machine learning pipelines

of IMU data for prediction of clinically relevant activities and biomechanics could help address this gap in the literature.

### 7.4.2 Individualised Movement Patterns can be Clinically Relevant

The authors of the OARSI recommendations for performance-based tests to assess physical function in people with knee osteoarthritis suggest that IMU technology could be used for both activity monitoring and motion analysis to complement patient-reported outcome measures and performance-based tests (Dobson et al., 2013). There are several studies demonstrating the clinical application IMUs for people with knee osteoarthritis (Kobsar & Ferber, 2018; Kobsar et al., 2017; Wang et al., 2021) and other health conditions (Dorsch et al., 2014; Ginis et al., 2016; Kent et al., 2015; Wernli, O'Sullivan, et al., 2020). Some studies have demonstrated a relationship between changes in movement parameters and clinical outcomes. For example, an experimental single-case series of low back pain demonstrated a relationship between a change in movement measured with IMUs, and change in activity limitation or pain occurred 74% of the time when the biomechanical outcomes were selected based on an activities reported in the Patient Specific Functional Scale (Wernli, O'Sullivan, et al., 2020). For the majority of the participants (10/12) in that study, there was a strong correlation ($r \geq 0.5$) between changes in movement pattern and clinical outcome for at least one biomechanical outcome. The authors of that study concluded that patients were demonstrating a 'less protective' movement pattern because of consistent changes across participants with increases in range of movement and speed with reduced muscle activation across a range of activities.

There is also early work demonstrating that IMUs data can be used to track participant-specific movement patterns in people with knee osteoarthritis (Kobsar & Ferber, 2018). That study used principal component analysis to reduce the high dimensionality of baseline accelerometer gait features, followed by a traditional machine learning approach (support vector machine) to determine if change occurred. Those results demonstrated that change in individualised principal component acceleration patterns from IMUs is highly correlated with patient-reported outcome measures (KOOS) (Spearman's rank correlation coefficient = 0.78, $r = 0.95$) (Kobsar

& Ferber, 2018). They used linear acceleration from three IMUs placed on the lower back, thigh, and shank to train and test the machine learning model. However, linear accelerations are not a typical movement parameter that have clear clinical relevance based on established evidence that could help guide clinical decisions, like those described in section 2.1.3.1 and 2.1.4. More traditionally, angular kinematics, joint forces or moments or muscle activity parameters have been used as measures of impairment or abnormal movement patterns. One issue recognised by Kobsar and Ferber (2018) is that univariate analysis of changes in commonly investigated movement parameters like knee adduction moment have questionable sensitivity, especially in group-based assessment for change – a consideration that is supported by the findings of the systematic review presented in Chapter 3. The authors describe the clinical application of their system would be to explain the percentage change or graphical representation of change for clinicians. That type of system could provide important information about change in acceleration data and change in clinical outcomes following an intervention. However, to our knowledge there are no studies that have identified if specific baseline acceleration characteristics are biomechanical risk factors like those studies that have investigated knee adduction moment. It is therefore unclear which *baseline* acceleration characteristics may be potential targets for clinical intervention to improve clinical outcomes or slow structural progression. In contrast, cross-sectional literature suggests that knee joint kinematic, kinetic and muscle activity parameters are related to clinical outcomes or structural progression which may have greater potential to inform clinicians about which movement patterns to target as part of a patient's overall management.

### 7.4.3 IMUs are Suited to Assessing Two Clinical Paradigms

IMUs have the potential to assist with clinical reasoning by collecting information that can be used by clinicians about the patient's baseline movement patterns and changes in movement patterns over time. Monitoring could be used to track progression or change in biomechanical status. Two clinical paradigms exist relevant to monitoring a patient using IMUs – (a) monitoring biomechanical risk factors for structural progression, and (b) assessment of avoidance behaviours and maladaptive movement patterns.

#### 7.4.3.1 Monitoring of Biomechanical Risk Factors

In the future, risk profiles could be identified that provide the opportunity for secondary prevention strategies (Mahmoudian et al., 2018) to be implemented that have the potential to slow the clinical and structural progression of knee osteoarthritis. Granted, it is currently unclear which patients respond to interventions that target movement parameters, and if those changes are causally linked to prevention of structural progression. But should research identify that some individuals be more responsive to a targeted change in movement patterns, part of a risk profile assessment in clinical practice could include IMU motion analysis in clinical and free-living environments.

For the activity of walking; knee adduction and flexion moments (Chehab et al., 2014) and less toe-out (Chang et al., 2007) are all risk factors for structural progression of medial knee osteoarthritis. Authors of those studies unanimously conclude that interventions should target these risk factors. Therefore, machine learning prediction models using IMU data, such as that presented in Chapter 4 and 6, provide an opportunity to assess and monitor risk factors like knee adduction moment.

Personal structural and biological risk factors also exist which could also be monitored, such as age, BMI, and genetic factors as part of a larger machine learning model for prediction of risk of structural progression (Cui et al., 2020; Silverwood et al., 2015). Other psychosocial (e.g. mood, fear, confidence) and lifestyle factors (e.g. sleep) associated with activity limitation and pain (Lentz et al., 2020) could also be monitored. As the cost of IMU systems continues to reduce, and artificial intelligence enters the mainstream of day-to-day clinical practice, clinicians may be provided with tools capable of recognising pre-clinical knee osteoarthritis and provide early intervention strategies targeted towards preventing or reducing structural progression. Alternatively, a clinician and patient may collectively wish to focus on monitoring activities or movement patterns for the purposes of addressing avoidance behaviours or maladaptive movement patterns that are related to activity limitation and pain.

#### 7.4.3.2 Avoidance Behaviours and Maladaptive Movement Patterns

Pain is thought to directly result in altered movement patterns (Hodges & Smeets, 2015; Hodges & Tucker, 2011). One alternative explanation for avoidance behaviours and altered movement patterns is that people with knee osteoarthritis adopt a 'careful mobility' strategy (Maly, 2009). Similar interpretations have been made across body regions describing 'protective' movement patterns (Hodges & Tucker, 2011; O'Sullivan et al., 2018). This interpretation is supported by both qualitative and quantitative studies. People with knee osteoarthritis describe adjusting their behaviour by avoiding or reducing pain provoking activities (Darlow et al., 2018; Maly & Krupa, 2007; Vlaeyen & Linton, 2000, 2012; Wallis et al., 2019). There is also evidence of fear and reduced confidence performing day-to-day activities (Caneiro et al., 2021; Maly & Krupa, 2007; Wallis et al., 2019), which can affect movement patterns (Hart, Collins, Ackland, Cowan, et al., 2015).

Together, changes in mobility and other behaviour in people with knee osteoarthritis can be interpreted as a means of protecting themselves from the experience of pain or of avoiding harm to their knee (Maly, 2009). Initially, behavioural responses to pain are considered to be adaptive, protective and helpful, but only in the short-term. Longer-term changes are thought to be maladaptive, provocative and unhelpful because they have a tendency to increase load, reduce movement and reduce movement variability during daily activities that collectively result in ongoing pain (Hodges & Tucker, 2011; O'Sullivan et al., 2018). Therefore, one potential intervention strategy for people with knee osteoarthritis, includes targeting maladaptive movement patterns during patient-specific clinically relevant activities for the purpose of symptom modification (Lehman, 2018; O'Sullivan et al., 2018; Preece et al., 2021).

### 7.4.4 Assessing Avoidance Behaviours and Maladaptive Movement Patterns

Patient-reported outcome measures, performance tests and monitoring of physical activity intensity do not provide objective data about the number of times or the length of time a person performs a range of clinically important activities in free-living environments. In qualitative studies, both patients and clinicians expressed the belief that IMU systems could be helpful for monitoring physical function,

motivating patients, and providing the clinician personalised information to assist with clinical decision-making (Papi et al., 2015; Papi et al., 2016). The machine learning models presented in this thesis provide a potential to objectively quantify IMU data to help assess activity avoidance behaviours or monitor biomechanical risk factors. For a clinician, having that information may help facilitate treatment selection or help provide motivating feedback to the patient. While there is evidence that IMUs in the form of an accelerometer provide information about physical activity intensity or active versus sedentary time (Frimpong et al., 2020), those approaches do not provide adequate richness of detail about which clinically relevant activities were performed, but rather the position or general movement of the body. For example, some wearable sensors like the activePAL can provide information about time spent sitting, standing, lying, and stepping, which are relevant activities for a person with knee osteoarthritis (Frimpong et al., 2020). However, that technology does not provide information about other clinically important activities such as whether a patient used the stairs or stood from a chair, nor directly provide biomechanical information about (potential) clinically relevant movement parameters.

An IMU-human activity recognition system like that presented in Chapter 4 has the potential to provide valuable information for a clinician about avoidance behaviours for specific activities (see section 7.4.4.1). Additionally, such a system could segment and label data for subsequent biomechanical processing as part of a data handling pipeline. It would then be theoretically possible to extract clinically useful biomechanical data from the segmented and labelled samples. Segmenting and labelling data would reduce the computational resources required and time spent by a clinician analysing many hours of biomechanical data.

During a consultation, a clinician may observe careful mobility strategies such as activity avoidance behaviours or maladaptive movement patterns and be interested in quantifying their observations using technology. Some clinicians have already adopted devices, such as smartphone biomechanical analysis applications and IMUs (based on fusion algorithms) into their assessments to quantify movement patterns, although those technologies have their limitations (see section 2.3.2.2 and 2.4.2). The machine learning prediction models using IMU data for kinematics (Chapter 5) and kinetics

(Chapter 6) provide a potential solution for the limitations that exist for current technology used in clinical practice. Beyond screening biomechanical risk factors, two other potential applications for IMUs include monitoring activity avoidance and maladaptive movement patterns.

#### 7.4.4.1 Monitoring Activity Avoidance

Careful mobility strategies may initially be adaptive to minimise acute pain. However, when pain is persistent, such behavioural responses may become maladaptive and result in reduced mobility, deconditioning and ongoing activity limitation and pain (Vlaeyen & Linton, 2000, 2012). One potential method of assessing avoidance behaviours could be to use IMUs record the amount of time or the frequency that a patient engages with an activity (Sparkes et al., 2019). Machine learning human activity monitoring using IMU data has the potential to provide information about the actual performance of clinically important activities in free-living environments. Two possible options for activity recognition are through the use of event detection algorithms (Fadillioglu et al., 2020), or machine learning based human activity recognition (Chapter 4). To provide data about activity avoidance, further development is required that integrates the human activity recognition machine learning model into a system that can provide information on time or frequency of activity engagement. In the future, human activity recognition could aid a clinician in their management of a patient by recording data that provides them objective information about avoidance behaviours. Combining patient-reported outcomes with human activity recognition based on IMU data could provide a clinician with a broader understanding of a patient's physical function and avoidance behaviours than patient-reported outcomes alone.

The following section is a hypothetical case scenario that draws together patient-specific information, clinical reasoning and how machine learning human activity recognition based on IMU data could be integrated in the management of a patient with knee osteoarthritis.

## CASE STUDY – NELLIE

Nellie developed insidious onset knee pain in her mid-50s. Now a decade later, she has recently experienced increasing symptoms over the past 6 months, resulting in her avoiding activities. She can walk as far as she likes and does not have pain when sitting or standing. Nellie tells her physiotherapist that she has difficulty with ascending staircases at work and home because of her knee osteoarthritis, resulting in avoidance behaviours that result in her using the elevator at work. At home Nellie goes upstairs, one step at a time, sideways, always leading with her pain free leg despite being capable of using a reciprocal gait pattern.

Her physiotherapist records outcome measures, including the Patient-Specific Functional Scale and Stair Climb Performance Test (Dobson et al., 2013). On the Patient-Specific Functional Scale, Nellie rates the difficulty using stairs as 8/10 at work, and she takes 22 seconds to ascend and descend a nine-step staircase.

In her physiotherapy consultations, she progresses well, building strength and confidence ascending a small staircase in the clinic with a reciprocal gait pattern. However, she still avoids stairs at work because she lacks confidence. The physiotherapist suggests that Nellie wear some IMUs to collect some information about how she is moving during the day. The clinician shows Nellie how to place the IMUs on her legs. She goes to work and completes a normal day. Data from the IMUs is uploaded to a cloud-based storage service, and automatically processed using the human activity recognition machine learning algorithm for people with knee osteoarthritis. No instances of ascending stairs are noted during work hours, with only one occasion before work and a few after work.

Both Nellie and the physiotherapist agree, while the performance in the clinic has improved, they need to try some

behavioural modifications (O'Sullivan et al., 2018; Preece et al., 2021) to build her confidence relating to stair use at work. One behavioural intervention includes setting up a reminder system to encourage Nellie to use the staircase. She agrees that a few days per week she will wear the IMUs for the purposes of using that data to count the number of times she uses the staircase each day.

The first day after commencing the behaviour change intervention while wearing the IMUs, Nellie used the staircase on two occasions. By the final workday and third time wearing the IMUs, Nellie used the stairs on five occasions. After four more weeks of wearing the IMUs, the IMU-based human activity recognition system indicates Nellie is using the staircase at work consistently more than eight times per day. She rates her difficulty ascending the staircase as 3/10 on the Patient Specific Functional Scale and now takes 15 seconds to ascend and descend nine stairs.

---

This case scenario demonstrates how different physical function outcomes can provide different information for a clinician and how outcomes can be individualised, in this case towards collecting information about the use of stairs. As part of that assessment of physical function, machine learning prediction models for human activity recognition can help provide a fuller picture of physical function for a clinician based on actual performance during a usual workday.

There are certainly aspects that a clinician would need to consider, such as the number of days required to establish a stable baseline, number of days worn to establish change has occurred, the patient's technological literacy, internet speeds for uploading data, the validation and statistical accuracy of the model (see section 7.3.4), and further work is required to develop a real-time system. However, the hypothetical case scenario above demonstrates the potential of human activity recognition to provide information to both the physiotherapist and patient about changes in how a person engages with activities performed in free-living environments.

#### 7.4.4.2 Monitoring Maladaptive Movement Patterns

As already discussed, careful mobility strategies (i.e. altered movement patterns) may be adaptive in the short term but maladaptive in the long term (Vlaeyen & Linton, 2000, 2012). While changes in movement patterns may not be related to improvement in activity or pain at a group-level (Chapter 3), there remains the possibility that individualised assessment and monitoring of movement patterns could help target management. There is preliminary evidence that targeting individualised maladaptive movement patterns for the purposes of symptom control in people with knee osteoarthritis is possible as part of a broader integrated biopsychosocial intervention (Preece et al., 2021).

Two common patterns that exist in the population of people with knee osteoarthritis when walking are a 'quadriceps avoidance pattern' (Al-Zahrani & Bakheit, 2002; Fisher et al., 1997; Messier et al., 1992) and a 'flexion loading pattern' (Childs et al., 2004; Hart, Collins, Ackland, Cowan, et al., 2015; Heiden et al., 2009). During the stance phase of walking, a person with a quadriceps avoidance pattern maintains the knee in relative extension, whereas a person with a flexion loading pattern maintains the knee in relative flexion. To monitor changes in these two types of patterns (or other individual patterns), a clinician could use a data from IMUs.

In the first part of the IMU data handling pipeline, a human activity recognition system could segment and label data of specific activity directions (e.g. stand-to-sit) or phases (e.g. stance phase of descending stairs) of interest. Those data could be subsequently handled by biomechanical estimation or prediction methods to provide information about an individual's movement parameters. Possible methods for biomechanical estimation include fusion algorithms for estimating kinematics (Weygers et al., 2020), inverse dynamics for estimating kinetics (Karatsidis et al., 2019) or alternatively, machine learning to predict both kinematics and kinetics (Chapters 5 and 6).

Together, IMU facilitated human activity recognition and biomechanical analysis could provide a more comprehensive understanding about improvements in a patient's maladaptive avoidance behaviours and movement behaviours. That information could help provide insight into the relationship between movement

patterns, activity limitation, and pain. The machine learning models based on IMU data have the potential to help individualise physical assessment of people with knee osteoarthritis for people seeking healthcare, as well as inform clinical research including relationships between movement patterns and clinical outcomes. There is potential for healthcare of the future to include telehealth and real time biofeedback applications outside of clinical environments by harnessing rapidly improving technology including improved internet speeds, cloud-based servers, optimised machine learning models and IMUs.

### 7.4.5 Summary

Clinical guidelines recommend individualised assessment and management for people with knee osteoarthritis. One approach to individualising assessment of movement patterns to support clinical decision making and facilitate individualised management is using IMUs. Machine learning data handling pipelines for IMU data have the potential to facilitate assessment of biomechanical risk factors for structural progression as well as monitor activity avoidance and maladaptive movement patterns during a clinical encounter or in free-living environments. With further development of machine learning models that use IMUs data, potentially combined with other improving technologies, there are opportunities for remote monitoring that may benefit telehealth and real time biofeedback.

## 7.5 Strengths and Limitations of Thesis

The systematic review in Chapter 3 has both methodological and clinically relevant strengths. Methodological strengths include prospective registration, use of PRISMA reporting guidelines and GRADE assessment. The search strategy was facilitated by a senior faculty librarian, and two independent reviewers. As the question was about a relationship, rather than efficacy, another strength is the inclusion of cohort studies and randomised controlled trials, and reporting on within-group change or correlation analysis rather than between-group mean difference. This review also has the additional benefit of including a wide range of movement parameters. However, because the overall quality of the included studies was low, most studies reported only within-group mean change, and the only activity that was investigated across studies was walking, the resultant confidence of our findings is limited. Additional limitations include risk of selective reporting bias (studies were not prospectively registered), language bias (only English) and publication bias. We also estimated means and standard deviations for studies that did not report those values at one time point which may have introduced some imprecision.

The machine learning studies have a range of strengths. Most notably, the work within Chapters 4 to 6 represents an interdisciplinary collaboration between clinicians, clinical researchers, data scientists, biomechanists, and biostatisticians using a data-centric approach to developing the machine learning models. We present in Chapters 4 to 6 the key components of an innovative data handling pipeline for the use of IMU technology for people with knee osteoarthritis. Each model's performance was comparable to that reported in previous studies that included healthy participants or those with other health conditions when using similar validation approaches. In Chapter 4, we present the first machine learning human activity recognition system validated on people with knee osteoarthritis. One important strength of the human activity recognition system was that it was designed with two key clinical ideas in mind, (a) identifying activities recommended in guidelines for assessment of physical function, and (b) identifying directions of movement or phases of activities that would be useful for subsequent biomechanical analysis. The primary strength of the kinematic prediction model was the inclusion of

multiple activities, rather than only walking, potentially increasing its generalisability. Similarly, a strength of the kinetic prediction model was the inclusion of multiple movement parameters, rather than only knee adduction moment. Those biomechanical prediction studies also lay the ground for future studies that might consider the use of double- or single-leg models or patient-specific model development.

The machine learning studies also have several limitations. The sample size for the proof-of-concept machine learning studies were based on previous studies (Arif et al., 2015; Findlow et al., 2008; Fridriksdottir et al., 2020; Hendry et al,. 2020; Stetter et al., 2020; Stetter et al., 2019; Wouda et al., 2018), but were not properly powered. Now that the studies within this thesis have established the prediction accuracy of deep learning models, further sufficiently powered studies are required. While the machine learning models have comparable accuracy to previous studies, improving that accuracy may be required for clinical implementation. However, the acceptable level of accuracy for clinical decision making is not yet clear and is likely to be context-specific, dependent on decision-making about clinical factors, machine learning architecture, and data collection approaches. Most notably, the participants included in the study were predominantly male and of normal BMI which is not representative of the broader population with knee osteoarthritis and potentially limits the generalisability of the models. As the models were trained and validated on data collected in a laboratory environment, it is not yet clear if they are valid for data collected outside of a laboratory environment, especially on differing chair or step heights. The models were developed as proof-of-concept and therefore included only a small number of participants which may limit generalisability. The proposed data handling pipeline for IMU data using the human activity recognition and biomechanical prediction models has yet to be tested and different machine learning architecture have not yet been tested in head-to-head comparisons. Participant-specific models developed in Chapter 6 require a reference standard to train the model, which is not available in clinical practice. Big data approaches such as training models from a database using a 'closest match' is an alternative approach.

Collectively, these limitations and others are discussed within earlier sections of this thesis alongside potential solutions.

## 7.6 Future Directions

The research in this thesis brought together concepts relevant to clinical practice, clinical research, and data science using IMU data and machine learning to facilitate individualised clinical assessment and management. The intersection of these concepts provides considerable opportunity for future research, and ideas about such future research have been integrated into each section of this discussion. There is interdependence between (a) exploring if it is clinically important to target a change in movement patterns and objectively quantify changes in movement parameters, and (b) developing data handling systems suitable to answer those questions of clinical importance that could also eventually be implemented in clinical practice.

IMUs have the potential to be able to collect individual person-level data, across multiple environments, allowing them to be used beyond a research laboratory. IMUs could be used to capture individual person-level movement data about activities identified on a patient-reported outcome measure such as the Patient-Specific Functional Scale. To replicate a clinical encounter or the patient's daily life, IMU data for such studies could be collected in clinical or free-living environments rather than a laboratory where people may move differently (Brodie et al., 2017; Brodie et al., 2016; Del Din et al., 2016; Dreischarf et al., 2016; Renggli et al., 2020; Robles-García et al., 2015; Weiss et al., 2011). A pipeline using machine learning models for both human activity recognition and biomechanical prediction could be developed and tested to output a range of movement parameters for each individual. Using a single-case experimental design using cross-correlations (Wernli, O'Sullivan, et al., 2020), or ideally a clinical trial using mediation analysis (Kent et al., 2019), an intervention could be introduced that targets a change in selected movement parameters for each individual based on each patient's symptom response (O'Sullivan et al., 2018; Preece et al., 2021).

Yet the proposed IMU system necessary to test the relationship between a change in movement parameters and clinical outcomes requires significant

development before implementation. Testing of the head-to-head performance of machine learning models is required for population-based generic models compared to individualised models (Ferrari et al., 2022). The current models need further validation for data collected in free-living environments. Large datasets (Ghorbani et al., 2021) or databases of IMU data from people with knee osteoarthritis may be required to be developed to build those models. Information contained within those databases ideally would include IMU data for multiple clinically important activities, levels of activity limitation and pain, as well as other biopsychosocial data that may influence the relationship between movement parameters and clinical outcomes (Dingenen et al., 2018). For the development of individualised models, pattern matching algorithms are needed that are designed to identify similar profiles across IMU and other biopsychosocial data to accommodate for heterogeneity across the population with knee osteoarthritis. While a database may help facilitate development of individualised models for human activity recognition allowing real-world assessment of activity avoidance behaviours, more nuanced systems that can provide information about the length of time or frequency a person performs an activity will be required. Data-centric approaches should also be considered to improve the performance of prediction models by optimising training data. For example, future studies could compare different methods for data collection, assessment of data quality, data cleaning, the effect of adding data from people with knee osteoarthritis to models designed for healthy people and explore the effect of adding simulated or augmented data to the training phase of model building (Mazumder et al., 2022; Ng, 2021).

## 7.7 Conclusions of Thesis

The relationship between movement patterns and activity limitation or pain in people with knee osteoarthritis is complex and challenging to assess, yet clinicians are tasked with determining the relevance of movement patterns in their clinical practice. The systematic review in Chapter 3 did not support the notion of a relationship between changes in movement patterns and changes in clinical outcomes. Several limitations existed in the current literature that preclude a clear understanding about a relationship at an individual person-level, despite guidelines clearly recommending

individualised assessment and management of people with knee osteoarthritis. Chapters 4 to 6 provide the groundwork for a possible solution to facilitate the individualised assessment of movement patterns using IMUs. With appropriate refinement of machine learning models there would potentially be an opportunity to explore the relationship between movement patterns and clinical outcomes while also addressing some of the limitations found within the current literature.

The machine learning models presented in this thesis were designed to address a gap in the research that is currently a significant barrier to implementation – that of how data collected in a free-living environment could potentially be handled to provide clinically relevant information. These models can handle IMU data in a way that is robust to electromagnetic interference and does not require calibration. The human activity recognition model presented in Chapter 4 provides a novel approach that could be useful for monitoring activity avoidance behaviours or segmenting data for subsequent biomechanical analysis. A data handling pipeline has been described that utilises machine learning models for human activity recognition followed by biomechanical prediction across various activities and movement parameters. With further work, the proposed pipeline could facilitate individualised assessment and monitoring of maladaptive movement patterns which may help clinical decision making. Establishing the performance of such a pipeline creates potential opportunities in both clinical practice and research for biomechanical risk prediction as well as telehealth and biofeedback.

Assessment of movement is only one part of the bigger clinical and data science picture. In the future, 'small data' from individualised biopsychosocial assessment (potentially including IMU data) may be leveraged to build large databases that could be used for clinical and research purposes. Personalised artificial intelligence models could be developed from 'big data' to facilitate clinical decision making or implement preventative measures that could help address the growing burden of knee osteoarthritis and other musculoskeletal health conditions.

## 7.8 Bibliography


Abbott, E., Campbell, A., Wise, E., Tidman, S. J., Lay, B. S., & Kent, P. (2022). Physiotherapists Could Detect Changes of 12 or More in Single-Plane Movement When Observing Forward Bending, Squat or Hand-over-Head: A Cross-Sectional Experiment. *Musculoskeletal Science and Practice*, 102594. https://doi.org/10.1016/j.msksp.2022.102594

Ackerman, I. N., Bohensky, M. A., Zomer, E., Tacey, M., Gorelik, A., Brand, C. A., & de Steiger, R. (2019). The Projected Burden of Primary Total Knee and Hip Replacement for Osteoarthritis in Australia to the Year 2030. *BMC Musculoskeletal Disorders*, *20*(1), 90. https://doi.org/10.1186/s12891-019-2411-9

Ackerman, I. N., Pratt, C., Gorelik, A., & Liew, D. (2018). Projected Burden of Osteoarthritis and Rheumatoid Arthritis in Australia: A Population-Level Analysis. *Arthritis Care & Research*, *70*(6), 877-883. https://doi.org/10.1002/acr.23414

Ageberg, E., & Roos, E. M. (2015). Neuromuscular Exercise as Treatment of Degenerative Knee Disease. *Exercise and Sport Sciences Reviews*, *43*(1), 14-22. https://doi.org/10.1249/jes.0000000000000030

Ahamed, N. U., Kobsar, D., Benson, L. C., Clermont, C. A., Osis, S. T., & Ferber, R. (2019). Subject-Specific and Group-Based Running Pattern Classification Using a Single Wearable Sensor. *Journal of Biomechanics*, *84*, 227-233. https://doi.org/10.1016/j.jbiomech.2019.01.001

Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine. *Database*, *2020*, baaa010. https://doi.org/10.1093/database/baaa010

Al-Khlaifat, L., Herrington, L. C., Hammond, A., Tyson, S. F., & Jones, R. K. (2016). The Effectiveness of an Exercise Programme on Knee Loading, Muscle Co-Contraction, and Pain in Patients with Medial Knee Osteoarthritis: A Pilot Study. *The Knee*, *23*(1), 63-69. https://doi.org/10.1016/j.knee.2015.03.014

Al-Zahrani, K., & Bakheit, A. (2002). A Study of the Gait Characteristics of Patients with Chronic Osteoarthritis of the Knee. *Disability and Rehabilitation*, *24*(5), 275-280. https://doi.org/10.1080/09638280110087098

Alberg, A. J., Park, J. W., Hager, B. W., Brock, M. V., & Diener-West, M. (2004). The Use of "Overall Accuracy" to Evaluate the Validity of Screening or Diagnostic Tests. *Journal of general internal medicine*, *19*(5 Pt 1), 460-465. https://doi.org/10.1111/j.1525-1497.2004.30091.x

Albert, M., Toledo, S., Shapiro, M., & Koerding, K. (2012). Using Mobile Phones for Activity Recognition in Parkinson's Patients. *Frontiers in Neurology*, *3*(158). https://doi.org/10.3389/fneur.2012.00158

Alentorn-Geli, E., Samuelsson, K., Musahl, V., Green, C. L., Bhandari, M., & Karlsson, J. (2017). The Association of Recreational and Competitive Running with Hip and Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Journal of Orthopaedic & Sports Physical Therapy*, *47*(6), 373-390. https://doi.org/10.2519/jospt.2017.7137

Alghadir, A. H., Anwer, S., Iqbal, A., & Iqbal, Z. A. (2018). Test-Retest Reliability, Validity, and Minimum Detectable Change of Visual Analog, Numerical

Rating, and Verbal Rating Scales for Measurement of Osteoarthritic Knee Pain. *Journal of Pain Research*, *11*, 851-856. https://doi.org/10.2147/JPR.S158847

Altman, R., Asch, E., Bloch, D., Bole, G., Borenstein, D., Brandt, K., . . . Wolfe, F. (1986). Development of Criteria for the Classification and Reporting of Osteoarthritis: Classification of Osteoarthritis of the Knee. *Arthritis & Rheumatism*, *29*(8), 1039-1049. https://doi.org/10.1002/art.1780290816

Andreu-Perez, J., Garcia-Gancedo, L., McKinnell, J., Van der Drift, A., Powell, A., Hamy, V., . . . Yang, G.-Z. (2017). Developing Fine-Grained Actigraphies for Rheumatoid Arthritis Patients from a Single Accelerometer Using Machine Learning. *Sensors*, *17*(9). https://doi.org/10.3390/s17092113

Apple Inc. (2022). *Workout Types on Apple Watch*. Retrieved October 24 from https://support.apple.com/en-au/HT207934

Argent, R., Drummond, S., Remus, A., O'Reilly, M., & Caulfield, B. (2019). Evaluating the Use of Machine Learning in the Assessment of Joint Angle Using a Single Inertial Sensor. *Journal of Rehabilitation and Assistive Technologies Engineering*, *6*, 2055668319868544. https://doi.org/10.1177%2F2055668319868544

Arif, M., & Kattan, A. (2015). Physical Activities Monitoring Using Wearable Acceleration Sensors Attached to the Body. *PLOS ONE*, *10*(7), e0130851. https://doi.org/10.1371/journal.pone.0130851

Arkorful, V. E., Shuliang, Z., Muhideen, S., Basiru, I., & Hammond, A. (2020). An Empirical Investigation of Health Practitioners Technology Adoption: The Mediating Role of Electronic Health. *International Journal of Public Administration*, *43*(12), 1013-1028. https://doi.org/10.1080/01900692.2019.1664569

Ascioglu, G., & Senol, Y. (2020). Design of a Wearable Wireless Multi-Sensor Monitoring System and Application for Activity Recognition Using Deep Learning. *IEEE Access*, *8*, 169183-169195. https://doi.org/10.1109/ACCESS.2020.3024003

Astephen, J. L., Deluzio, K. J., Caldwell, G. E., Dunbar, M. J., & Hubley-Kozey, C. L. (2008). Gait and Neuromuscular Pattern Changes Are Associated with Differences in Knee Osteoarthritis Severity Levels. *Journal of Biomechanics*, *41*(4), 868-876. https://doi.org/10.1016/j.jbiomech.2007.10.016

Australian Institute of Health and Welfare. (2021). *Osteoarthritis*. https://www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/osteoarthritis/contents/what-is-osteoarthritis

Bachmann, E. R., Yun, X., & Peterson, C. W. (2004). An Investigation of the Effects of Magnetic Variations on Inertial/Magnetic Orientation Sensors. IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004,

Baliunas, A. J., Hurwitz, D. E., Ryals, A. B., Karrar, A., Case, J. P., Block, J. A., & Andriacchi, T. P. (2002). Increased Knee Joint Loads During Walking Are Present in Subjects with Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *10*(7), 573-579. https://doi.org/10.1053/joca.2002.0797

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., . . . Norris, S. (2011). Grade Guidelines: 3. Rating the Quality of Evidence. *Journal of Clinical Epidemiology*, *64*(4), 401-406. https://doi.org/10.1016/j.jclinepi.2010.07.015

Bannuru, R. R., Osani, M. C., Vaysbrot, E. E., Arden, N. K., Bennell, K., Bierma-Zeinstra, S. M. A., . . . McAlindon, T. E. (2019). Oarsi Guidelines for the Non-Surgical Management of Knee, Hip, and Polyarticular Osteoarthritis. *Osteoarthritis and Cartilage*, *27*(11), 1578-1589. https://doi.org/10.1016/j.joca.2019.06.011

Bartlett, R. (2007). *Introduction to Sports Biomechanics: Analysing Human Movement Patterns*. Routledge.

Bellamy, N., Buchanan, W. W., Goldsmith, C. H., Campbell, J., & Stitt, L. W. (1988). Validation Study of Womac: A Health Status Instrument for Measuring Clinically Important Patient Relevant Outcomes to Antirheumatic Drug Therapy in Patients with Osteoarthritis of the Hip or Knee. *The Journal of Rheumatology*, *15*(12), 1833-1840. https://psycnet.apa.org/record/2019-29410-001

Bennell, K. L., Bowles, K. A., Wang, Y., Cicuttini, F., Davies-Tuck, M., & Hinman, R. S. (2011). Higher Dynamic Medial Knee Load Predicts Greater Cartilage Loss over 12 Months in Medial Knee Osteoarthritis. *Annals of the Rheumatic Diseases*, *70*(10), 1770-1774. https://doi.org/10.1136/ard.2010.147082

Bennell, K. L., Dobson, F., Roos, E. M., Skou, S. T., Hodges, P., Wrigley, T. V., . . . Hinman, R. S. (2015). Influence of Biomechanical Characteristics on Pain and Function Outcomes from Exercise in Medial Knee Osteoarthritis and Varus Malalignment: Exploratory Analyses from a Randomized Controlled Trial. *Arthritis Care & Research*, *67*(9), 1281-1288. https://doi.org/10.1002/acr.22558

Bennell, K. L., Hunt, M. A., Wrigley, T. V., Hunter, D. J., McManus, F. J., Hodges, P. W., . . . Hinman, R. S. (2010). Hip Strengthening Reduces Symptoms but Not Knee Load in People with Medial Knee Osteoarthritis and Varus Malalignment: A Randomised Controlled Trial. *Osteoarthritis and Cartilage*, *18*(5), 621-628. https://doi.org/10.1016/j.joca.2010.01.010

Bennell, K. L., Kyriakides, M., Metcalf, B., Egerton, T., Wrigley, T. V., Hodges, P. W., . . . Hinman, R. S. (2014). Neuromuscular Versus Quadriceps Strengthening Exercise in Patients with Medial Knee Osteoarthritis and Varus Malalignment: A Randomized Controlled Trial. *Arthritis & Rheumatology*, *66*(4), 950-959. https://doi.org/10.1002/art.38317

Bensalma, F., Hagemeister, N., Cagnin, A., Ouakrim, Y., Bureau, N. J., Choinière, M., & Mezghani, N. (2022). Biomechanical Markers Associations with Pain, Symptoms, and Disability Compared to Radiographic Severity in Knee Osteoarthritis Patients: A Secondary Analysis from a Cluster Randomized Controlled Trial. *BMC Musculoskeletal Disorders*, *23*(1), 896. https://doi.org/10.1186/s12891-022-05845-1

Bensalma, F., Mezghani, N., Ouakrim, Y., Fuentes, A., Choinière, M., Bureau, N. J., . . . Hagemeister, N. (2019). A Multivariate Relationship between the Kinematic and Clinical Parameters of Knee Osteoarthritis Population. *BioMedical Engineering OnLine*, *18*(1), 58. https://doi.org/10.1186/s12938-019-0676-8

Besier, T. F., Sturnieks, D. L., Alderson, J. A., & Lloyd, D. G. (2003). Repeatability of Gait Data Using a Functional Hip Joint Centre and a Mean Helical Knee Axis. *Journal of Biomechanics*, *36*(8), 1159-1168. https://doi.org/10.1016/S0021-9290(03)00087-3

Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). What Is Machine Learning? A Primer for the Epidemiologist. *American Journal of Epidemiology*, *188*(12), 2222-2239. https://doi.org/10.1093/aje/kwz189

Binnie, T., O'Sullivan, P., Bunzli, S., Campbell, A., Ng, L., Kent, P., & Smith, A. (2022). How Do People with Knee Osteoarthritis Conceptualize Knee Confidence? A Qualitative Study. *Physical Therapy*, pzac082. 10.1093/ptj/pzac082

Binnie, T., Smith, A., Kent, P., Ng, L., O'Sullivan, P., Tan, J.-S., . . . Campbell, A. (2021). Concurrent Validation of Inertial Sensors for Measurement of Knee Kinematics in Individuals with Knee Osteoarthritis: A Technical Report. *Health and Technology*, 1-10. https://doi.org/10.1007/s12553-021-00616-9

Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical Practice as Natural Laboratory for Psychotherapy Research: A Guide to Case-Based Time-Series Analysis. *American psychologist*, *63*(2), 77. https://doi.org/10.1037/0003-066X.63.2.77

Bouchouras, G., Patsika, G., Hatzitaki, V., & Kellis, E. (2015). Kinematics and Knee Muscle Activation During Sit-to-Stand Movement in Women with Knee Osteoarthritis. *Clinical Biomechanics*, *30*(6), 599-607. https://doi.org/10.1016/j.clinbiomech.2015.03.025

Brenneman, E. C., Kuntz, A. B., Wiebenga, E. G., & Maly, M. R. (2015). A Yoga Strengthening Program Designed to Minimize the Knee Adduction Moment for Women with Knee Osteoarthritis: A Proof-of-Principle Cohort Study. *PLOS ONE*, *10*(9), e0136854. https://doi.org/10.1371/journal.pone.0136854

Briani, R. V., Ferreira, A. S., Pazzinatto, M. F., Pappas, E., De Oliveira Silva, D., & Azevedo, F. M. d. (2018). What Interventions Can Improve Quality of Life or Psychosocial Factors of Individuals with Knee Osteoarthritis? A Systematic Review with Meta-Analysis of Primary Outcomes from Randomised Controlled Trials. *British Journal of Sports Medicine*, *52*(16), 1031. https://doi.org/10.1136/bjsports-2017-098099

Brock, H., Ohgi, Y., & Lee, J. (2017). Learning to Judge Like a Human: Convolutional Networks for Classification of Ski Jumping Errors. Proceedings of the 2017 ACM international symposium on wearable computers,

Brodie, M. A., Coppens, M. J., Ejupi, A., Gschwind, Y. J., Annegarn, J., Schoene, D., . . . Delbaere, K. (2017). Comparison between Clinical Gait and Daily-Life Gait Assessments of Fall Risk in Older People. *Geriatrics & Gerontology International*, *17*(11), 2274-2282. https://doi.org/10.1111/ggi.12979

Brodie, M. A., Coppens, M. J. M., Lord, S. R., Lovell, N. H., Gschwind, Y. J., Redmond, S. J., . . . Delbaere, K. (2016). Wearable Pendant Device Monitoring Using New Wavelet-Based Methods Shows Daily Life and Laboratory Gaits Are Different. *Medical & Biological Engineering & Computing*, *54*(4), 663-674. 10.1007/s11517-015-1357-9

Brouwer, G. M., Tol, A. W. V., Bergink, A. P., Belo, J. N., Bernsen, R. M. D., Reijman, M., . . . Bierma-Zeinstra, S. M. A. (2007). Association between Valgus and Varus Alignment and the Development and Progression of Radiographic Osteoarthritis of the Knee. *Arthritis & Rheumatism*, *56*(4), 1204-1211. https://doi.org/10.1002/art.22515

Bulling, A., Blanke, U., & Schiele, B. (2014). A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors. *ACM Computing Surveys (CSUR)*, *46*(3), 1-33. https://doi.org/10.1145/2499621

Caballé-Cervigón, N., Castillo-Sequera, J. L., Gómez-Pulido, J. A., Gómez-Pulido, J. M., & Polo-Luque, M. L. (2020). Machine Learning Applied to Diagnosis of Human Diseases: A Systematic Review. *Applied Sciences*, *10*(15 ), 5135. https://doi.org/10.3390/app10155135

Calvet, J., Orellana, C., Larrosa, M., Navarro, N., Chillarón, J. J., Pedro-Botet, J., . . . Gratacós, J. (2016). High Prevalence of Cardiovascular Co-Morbidities in Patients with Symptomatic Knee or Hand Osteoarthritis. *Scandinavian Journal of Rheumatology*, *45*(1), 41-44. https://doi.org/10.3109/03009742.2015.1054875

Camomilla, V., Cereatti, A., Cutti, A. G., Fantozzi, S., Stagni, R., & Vannozzi, G. (2017). Methodological Factors Affecting Joint Moments Estimation in Clinical Gait Analysis: A Systematic Review. *BioMedical Engineering OnLine*, *16*(1), 106. https://doi.org/10.1186/s12938-017-0396-x

Caneiro, J. P., O'Sullivan, P. B., Roos, E. M., Smith, A. J., Choong, P., Dowsey, M., . . . Barton, C. J. (2020). Three Steps to Changing the Narrative About Knee Osteoarthritis Care: A Call to Action. *British Journal of Sports Medicine*, *54*(5), 256. http://doi.org/10.1136/bjsports-2019-101328

Caneiro, J. P., Roos, E. M., Barton, C. J., Sullivan, K., Kent, P., Lin, I., . . . Sullivan, P. (2020). It Is Time to Move Beyond 'Body Region Silos' to Manage Musculoskeletal Pain: Five Actions to Change Clinical Practice. *British Journal of Sports Medicine*, *54*(8), 438. https://doi.org/10.1136/bjsports-2018-100488

Caneiro, J. P., Smith, A., Bunzli, S., Linton, S., Moseley, G. L., & O'Sullivan, P. (2021). From Fear to Safety: A Roadmap to Recovery from Musculoskeletal Pain. *Physical Therapy*, pzab271. https://doi.org/10.1093/ptj/pzab271

Canetti, E. F. D., Schram, B., Orr, R. M., Knapik, J., & Pope, R. (2020). Risk Factors for Development of Lower Limb Osteoarthritis in Physically Demanding Occupations: A Systematic Review and Meta-Analysis. *Applied Ergonomics*, *86*, 103097. https://doi.org/10.1016/j.apergo.2020.103097

Carpenter, C., Yang, C.-H., & West, D. (2021). A Comparison of Sedentary Behavior as Measured by the Fitbit and Activpal in College Students. *International Journal of Environmental Research and Public Health*, *18*(8). https://doi.org/10.3390/ijerph18083914

Chang, A., Almagor, O., Chmiel, J., Moisio, K., Zhang, J., Cahue, S., . . . Sharma, L. (2011). Frontal Plane Motion During Gait and Knee Confidence in Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *19*, S87-S88. https://doi.org/10.1016/S1063-4584(11)60202-1

Chang, A., Hurwitz, D., Dunlop, D., Song, J., Cahue, S., Hayes, K., & Sharma, L. (2007). The Relationship between Toe-out Angle During Gait and Progression of Medial Tibiofemoral Osteoarthritis. *Annals of the Rheumatic Diseases*, *66*(10), 1271. https://doi.org/10.1136/ard.2006.062927

Chang, A. H., Moisio, K. C., Chmiel, J. S., Eckstein, F., Guermazi, A., Prasad, P. V., . . . Sharma, L. (2015). External Knee Adduction and Flexion Moments During Gait and Medial Tibiofemoral Disease Progression in Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *23*(7), 1099-1106. https://doi.org/10.1016/j.joca.2015.02.005

Chapple, C. M., Nicholson, H., Baxter, G. D., & Abbott, J. H. (2011). Patient Characteristics That Predict Progression of Knee Osteoarthritis: A Systematic Review of Prognostic Studies. *Arthritis Care & Research*, *63*(8), 1115-1125. https://doi.org/10.1002/acr.20492

Charlton, P. C., Kenneally-Dabrowski, C., Sheppard, J., & Spratford, W. (2017). A Simple Method for Quantifying Jump Loads in Volleyball Athletes. *Journal of Science and Medicine in Sport*, *20*(3), 241-245. https://doi.org/10.1016/j.jsams.2016.07.007

Chehab, E. F., Favre, J., Erhart-Hledik, J. C., & Andriacchi, T. P. (2014). Baseline Knee Adduction and Flexion Moments During Walking Are Both Associated with 5 year Cartilage Changes in Patients with Medial Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *22*(11), 1833-1839. https://doi.org/10.1016/j.joca.2014.08.009

Chen, J.-L., Dai, Y.-N., Grimaldi, N. S., Lin, J.-J., Hu, B.-Y., Wu, Y.-F., & Gao, S. (2022). Plantar Pressure-Based Insole Gait Monitoring Techniques for Diseases Monitoring and Analysis: A Review. *Advanced Materials Technologies*, *7*(1), 2100566. https://doi.org/10.1002/admt.202100566

Chen, K.-H., Chen, P.-C., Liu, K.-C., & Chan, C.-T. (2015). Wearable Sensor-Based Rehabilitation Exercise Assessment for Knee Osteoarthritis. *Sensors*, *15*(2). https://doi.org/10.3390/s150204193

Chen, Y., & Xue, Y. (2015). A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer. 2015 IEEE international conference on systems, man, and cybernetics,

Cheung, R. T. H., Ho, K. K. W., Au, I. P. H., An, W. W., Zhang, J. H. W., Chan, Z. Y. S., . . . Rainbow, M. J. (2018). Immediate and Short-Term Effects of Gait Retraining on the Knee Joint Moments and Symptoms in Patients with Early Tibiofemoral Joint Osteoarthritis: A Randomized Controlled Trial. *Osteoarthritis and Cartilage*, *26*(11), 1479-1486. https://doi.org/10.1016/j.joca.2018.07.011

Childs, J. D., Sparto, P. J., Fitzgerald, G. K., Bizzini, M., & Irrgang, J. J. (2004). Alterations in Lower Extremity Movement and Muscle Activation Patterns in Individuals with Knee Osteoarthritis. *Clinical Biomechanics*, *19*(1), 44-49. https://doi.org/10.1016/j.clinbiomech.2003.08.007

Cisternas, M. G., Murphy, L., Sacks, J. J., Solomon, D. H., Pasta, D. J., & Helmick, C. G. (2016). Alternative Methods for Defining Osteoarthritis and the Impact on Estimating Prevalence in a Us Population-Based Survey. *Arthritis Care & Research*, *68*(5), 574-580. https://doi.org/10.1002/acr.22721

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Averaging Correlations: Expected Values and Bias in Combined Pearson Rs and Fisher's Z Transformations. *The Journal of General Psychology*, *125*(3), 245-261. https://doi.org/10.1080/00221309809595548

Coskun, D., Incel, O. D., & Ozgovde, A. (2015). Phone Position/Placement Detection Using Accelerometer: Impact on Activity Recognition. 2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP),

Cudejko, T., Button, K., Willott, J., & Al-Amri, M. (2021). Applications of Wearable Technology in a Real-Life Setting in People with Knee Osteoarthritis: A

Systematic Scoping Review. *Journal of Clinical Medicine*, *10*(23), 5645. https://doi.org/10.3390/jcm10235645

Cuesta-Vargas, A. I., Galán-Mercant, A., & Williams, J. M. (2010). The Use of Inertial Sensors System for Human Motion Analysis. *Physical Therapy Reviews*, *15*(6), 462-473. https://doi.org/10.1179/1743288X11Y.0000000006

Cui, A., Li, H., Wang, D., Zhong, J., Chen, Y., & Lu, H. (2020). Global, Regional Prevalence, Incidence and Risk Factors of Knee Osteoarthritis in Population-Based Studies. *EClinicalMedicine*, *29*. https://doi.org/10.1016/j.eclinm.2020.100587

Culvenor, A. G., Ruhdorfer, A., Juhl, C., Eckstein, F., & Øiestad, B. E. (2017). Knee Extensor Strength and Risk of Structural, Symptomatic, and Functional Decline in Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Arthritis Care & Research*, *69*(5), 649-658. https://doi.org/10.1002/acr.23005

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and Deep Learning for Sport-Specific Movement Recognition: A Systematic Review of Model Development and Performance. *Journal of Sports Sciences*, *37*(5), 568-600. https://doi.org/10.1080/02640414.2018.1521769

D'Souza, N., Charlton, J., Grayson, J., Kobayashi, S., Hutchison, L., Hunt, M., & Simic, M. (2022). Are Biomechanics During Gait Associated with the Structural Disease Onset and Progression of Lower Limb Osteoarthritis? A Systematic Review and Meta-Analysis. *Osteoarthritis and Cartilage*, *30*(3), 381-394. https://doi.org/10.1016/j.joca.2021.10.010

da Silva, F. S., de Melo, F. E. S., do Amaral, M. M. G., Caldas, V. V. A., Pinheiro, I. L. D., Abreu, B. J., & Vieira, W. H. B. (2015). Efficacy of Simple Integrated Group Rehabilitation Program for Patients with Knee Osteoarthritis: Single-Blind Randomized Controlled Trial. *Journal of Rehabilitation Research & Development*, *52*, 309+. https://doi.org/10.1682/JRRD.2014.08.0199

Darlow, B., Brown, M., Thompson, B., Hudson, B., Grainger, R., McKinlay, E., & Abbott, J. H. (2018). Living with Osteoarthritis Is a Balancing Act: An Exploration of Patients' Beliefs About Knee Pain. *BMC Rheumatology*, *2*(1), 15. https://doi.org/10.1186/s41927-018-0023-x

Davis, A. M., Davis, K. D., Skou, S. T., & Roos, E. M. (2020). Why Is Exercise Effective in Reducing Pain in People with Osteoarthritis? *Current Treatment Options in Rheumatology*, *6*(3), 146-159. https://doi.org/10.1007/s40674-020-00154-x

Davis, H. C., Luc-Harkey, B. A., Seeley, M. K., Troy Blackburn, J., & Pietrosimone, B. (2019). Sagittal Plane Walking Biomechanics in Individuals with Knee Osteoarthritis after Quadriceps Strengthening. *Osteoarthritis and Cartilage*, *27*(5), 771-780. https://doi.org/10.1016/j.joca.2018.12.026

de Rooij, M., van der Leeden, M., Heymans, M. W., Holla, J. F. M., Häkkinen, A., Lems, W. F., . . . Dekker, J. (2016). Prognosis of Pain and Physical Functioning in Patients with Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Arthritis Care & Research*, *68*(4), 481-492. https://doi.org/10.1002/acr.22693

de Vries, W. H. K., Veeger, H. E. J., Baten, C. T. M., & van der Helm, F. C. T. (2009). Magnetic Distortion in Motion Labs, Implications for Validating Inertial Magnetic Sensors. *Gait & Posture*, *29*(4), 535-541. https://doi.org/10.1016/j.gaitpost.2008.12.004

DeBerardinis, J., Dufek, J. S., Trabia, M. B., & Lidstone, D. E. (2018). Assessing the Validity of Pressure-Measuring Insoles in Quantifying Gait Variables. *Journal of Rehabilitation and Assistive Technologies Engineering*, *5*, 2055668317752088. https://doi.org/10.1177/2055668317752088

Deep, S., & Zheng, X. (2019). Leveraging Cnn and Transfer Learning for Vision-Based Human Activity Recognition. 2019 29th International Telecommunication Networks and Applications Conference (ITNAC),

Dekker, J., van Dijk, G. M., & Veenhof, C. (2009). Risk Factors for Functional Decline in Osteoarthritis of the Hip or Knee. *Current Opinion in Rheumatology*, *21*(5). https://doi.org/10.1097/BOR.0b013e32832e6eaa

Del Din, S., Godfrey, A., Galna, B., Lord, S., & Rochester, L. (2016). Free-Living Gait Characteristics in Ageing and Parkinson's Disease: Impact of Environment and Ambulatory Bout Length. *Journal of NeuroEngineering and Rehabilitation*, *13*(1), 46. https://doi.org/10.1186/s12984-016-0154-5

Dell'Isola, A., Allan, R., Smith, S. L., Marreiros, S. S. P., & Steultjens, M. (2016). Identification of Clinical Phenotypes in Knee Osteoarthritis: A Systematic Review of the Literature. *BMC Musculoskeletal Disorders*, *17*(1), 425. https://doi.org/10.1186/s12891-016-1286-2

Dell'Isola, A., & Steultjens, M. (2018). Classification of Patients with Knee Osteoarthritis in Clinical Phenotypes: Data from the Osteoarthritis Initiative. *PLOS ONE*, *13*(1), e0191045. 10.1371/journal.pone.0191045

Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, *132*(20), 1920-1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

DeVita, P., Aaboe, J., Bartholdy, C., Leonardis, J. M., Bliddal, H., & Henriksen, M. (2018). Quadriceps-Strengthening Exercise and Quadriceps and Knee Biomechanics During Walking in Knee Osteoarthritis: A Two-Centre Randomized Controlled Trial. *Clinical Biomechanics*, *59*, 199-206. https://doi.org/10.1016/j.clinbiomech.2018.09.016

Dingenen, B., Blandford, L., Comerford, M., Staes, F., & Mottram, S. (2018). The Assessment of Movement Health in Clinical Practice: A Multidimensional Perspective. *Physical Therapy in Sport*, *32*, 282-292. https://doi.org/10.1016/j.ptsp.2018.04.008

Dobkin, B. H. (2013). Wearable Motion Sensors to Continuously Measure Real-World Physical Activities. *Current Opinion in Neurology*, *26*(6). https://journals.lww.com/co-neurology/Fulltext/2013/12000/Wearable_motion_sensors_to_continuously_measure.3.aspx

Dobson, F., Hinman, R. S., Hall, M., Terwee, C., Roos, E. M., & Bennell, K. L. (2012). Measurement Properties of Performance-Based Measures to Assess Physical Function in Hip and Knee Osteoarthritis: A Systematic Review. *Osteoarthritis and Cartilage*, *20*(12), 1548-1562. https://doi.org/10.1016/j.joca.2012.08.015

Dobson, F., Hinman, R. S., Roos, E. M., Abbott, J. H., Stratford, P., Davis, A. M., . . . Thumboo, J. (2013). Oarsi Recommended Performance-Based Tests to Assess Physical Function in People Diagnosed with Hip or Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *21*(8), 1042-1052. https://doi.org/10.1016/j.joca.2013.05.002

Dorsch, A. K., Thomas, S., Xu, X., Kaiser, W., & Dobkin, B. H. (2014). Sirract: An International Randomized Clinical Trial of Activity Feedback During Inpatient Stroke Rehabilitation Enabled by Wireless Sensing. *Neurorehabilitation and Neural Repair*, *29*(5), 407-415. https://doi.org/10.1177/1545968314550369

Dorschky, E., Nitschke, M., Martindale, C. F., van den Bogert, A. J., Koelewijn, A. D., & Eskofier, B. M. (2020). Cnn-Based Estimation of Sagittal Plane Walking and Running Biomechanics from Measured and Simulated Inertial Sensor Data. *Frontiers in Bioengineering and Biotechnology*, *8*. https://doi.org/10.3389/fbioe.2020.00604

Drapeaux, A., & Carlson, K. (2020). A Comparison of Inertial Motion Capture Systems: Dorsavi and Xsens. *2020*, *8*(3), 4. http://doi.org/10.7575/aiac.ijkss.v.8n.3p.24

Dreischarf, M., Pries, E., Bashkuev, M., Putzier, M., & Schmidt, H. (2016). Differences between Clinical "Snap-Shot" and "Real-Life" Assessments of Lumbar Spine Alignment and Motion – What Is the "Real" Lumbar Lordosis of a Human Being? *Journal of Biomechanics*, *49*(5), 638-644. https://doi.org/10.1016/j.jbiomech.2016.01.032

Dunlop, D., Song, J., Semanik, P., Sharma, L., & Chang, R. (2011). Physical Activity Levels and Functional Performance in the Osteoarthritis Initiative: A Graded Relationship. *Arthritis & Rheumatism*, *63*(1), 127-136. https://doi.org/10.1002/art.27760

Ehara, Y., Fujimoto, H., Miyazaki, S., Tanaka, S., & Yamamoto, S. (1995). Comparison of the Performance of 3d Camera Systems. *Gait & Posture*, *3*(3), 166-169. https://doi.org/10.1016/0966-6362(95)99067-U

Emmerzaal, J., De Brabandere, A., Vanrompay, Y., Vranken, J., Storms, V., De Baets, L., . . . Timmermans, A. (2020). Towards the Monitoring of Functional Status in a Free-Living Environment for People with Hip or Knee Osteoarthritis: Design and Evaluation of the Jolo Blended Care App. *Sensors*, *20*(23), 6967. https://doi.org/10.3390/s20236967

Epifanio, I., Ávila, C., Page, Á., & Atienza, C. (2008). Analysis of Multiple Waveforms by Means of Functional Principal Component Analysis: Normal Versus Pathological Patterns in Sit-to-Stand Movement. *Medical & Biological Engineering & Computing*, *46*(6), 551-561. https://doi.org/10.1007/s11517-008-0339-6

Fadillioglu, C., Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2020). Automated Gait Event Detection for a Variety of Locomotion Tasks Using a Novel Gyroscope-Based Algorithm. *Gait & Posture*, *81*, 102-108. https://doi.org/10.1016/j.gaitpost.2020.06.019

Fantini Pagani, C., Funken, J., Heinrich, K., Ellermann, A., Schmidt-Wiethoff, R., & Potthast, W. (2020). Predicting the Knee Adduction Moment after High Tibial Osteotomy in Patients with Medial Knee Osteoarthritis Using Dynamic Simulations. *The Knee*, *27*(1), 61-70. https://doi.org/10.1016/j.knee.2019.08.003

Favre, J., Erhart-Hledik, J. C., & Andriacchi, T. P. (2014). Age-Related Differences in Sagittal-Plane Knee Function at Heel-Strike of Walking Are Increased in Osteoarthritic Patients. *Osteoarthritis and Cartilage*, *22*(3), 464-471. https://doi.org/10.1016/j.joca.2013.12.014

Felson, D. T. (2010). Identifying Different Osteoarthritis Phenotypes through Epidemiology. *Osteoarthritis and Cartilage*, *18*(5), 601-604. https://doi.org/10.1016/j.joca.2010.01.007

Ferrari, A., Micucci, D., Mobilio, M., & Napoletano, P. (2022). Deep Learning and Model Personalization in Sensor-Based Human Activity Recognition. *Journal of Reliable Intelligent Environments*. https://doi.org/10.1007/s40860-021-00167-w

Ferreira, G. E., Robinson, C. C., Wiebusch, M., Viero, C. C. d. M., da Rosa, L. H. T., & Silva, M. F. (2015). The Effect of Exercise Therapy on Knee Adduction Moment in Individuals with Knee Osteoarthritis: A Systematic Review. *Clinical Biomechanics*, *30*(6), 521-527. https://doi.org/10.1016/j.clinbiomech.2015.03.028

Findlow, A., Goulermas, J. Y., Nester, C., Howard, D., & Kenney, L. P. J. (2008). Predicting Lower Limb Joint Kinematics Using Wearable Motion Sensors. *Gait & Posture*, *28*(1), 120-126. https://doi.org/10.1016/j.gaitpost.2007.11.001

Fisher, N. M., Gresham, G. E., Abrams, M., Hicks, J., Horrigan, D., & Pendergast, D. R. (1993). Quantitative Effects of Physical Therapy on Muscular and Functional Performance in Subjects with Osteoarthritis of the Knees. *Archives of Physical Medicine and Rehabilitation*, *74*(8), 840-847.

Fisher, N. M., White, S. C., Yack, H. J., Smolinski, R. J., & Pendergast, D. R. (1997). Muscle Function and Gait in Patients with Knee Osteoarthritis before and after Muscle Rehabilitation. *Disability and Rehabilitation*, *19*(2), 47-55. https://doi.org/10.3109/09638289709166827

Foroughi, N., Smith, R. M., Lange, A. K., Baker, M. K., Fiatarone Singh, M. A., & Vanwanseele, B. (2011a). Lower Limb Muscle Strengthening Does Not Change Frontal Plane Moments in Women with Knee Osteoarthritis: A Randomized Controlled Trial. *Clinical Biomechanics*, *26*(2), 167-174. https://doi.org/10.1016/j.clinbiomech.2010.08.011

Foroughi, N., Smith, R. M., Lange, A. K., Singh, M. A. F., & Vanwanseele, B. (2011). Progressive Resistance Training and Dynamic Alignment in Osteoarthritis: A Single-Blind Randomised Controlled Trial. *Clinical Biomechanics*, *26*(1), 71-77. https://doi.org/10.1016/j.clinbiomech.2010.08.013

Fransen, M., McConnell, S., Harmer, A. R., Van der Esch, M., Simic, M., & Bennell, K. L. (2015). Exercise for Osteoarthritis of the Knee: A Cochrane Systematic Review. *British Journal of Sports Medicine*, *49*(24), 1554-1557. http://dx.doi.org/10.1136/bjsports-2015-095424

Fridriksdottir, E., & Bonomi, A. G. (2020). Accelerometer-Based Human Activity Recognition for Patient Monitoring Using a Deep Neural Network. *Sensors*, *20*(22), 6424. https://doi.org/10.3390/s20226424

Frimpong, E., van der Jagt, D. R., Mokete, L., Pietrzak, J., Kaoje, Y. S., Smith, A., . . . Meiring, R. M. (2020). Improvements in Objectively Measured Activity Behaviors Do Not Correlate with Improvements in Patient-Reported Outcome Measures Following Total Knee Arthroplasty. *The Journal of Arthroplasty*, *35*(3), 712-719.e714. https://doi.org/10.1016/j.arth.2019.10.016

Fukutani, N., Iijima, H., Aoyama, T., Yamamoto, Y., Hiraoka, M., Miyanobu, K., . . . Matsuda, S. (2016). Knee Pain During Activities of Daily Living and Its Relationship with Physical Activity in Patients with Early and Severe Knee

Osteoarthritis. *Clinical Rheumatology*, *35*(9), 2307-2316. https://doi.org/10.1007/s10067-016-3251-8

Gaudreault, N., Mezghani, N., Turcot, K., Hagemeister, N., Boivin, K., & de Guise, J. A. (2011). Effects of Physiotherapy Treatment on Knee Osteoarthritis Gait Data Using Principal Component Analysis. *Clinical Biomechanics*, *26*(3), 284-291. https://doi.org/10.1016/j.clinbiomech.2010.10.004

Georgiev, T., & Angelov, A. K. (2019). Modifiable Risk Factors in Knee Osteoarthritis: Treatment Implications. *Rheumatology International*, *39*(7), 1145-1157. 10.1007/s00296-019-04290-z

Gholamiangonabadi, D., Kiselov, N., & Grolinger, K. (2020). Deep Neural Networks for Human Activity Recognition with Wearable Sensors: Leave-One-Subject-out Cross-Validation for Model Selection. *IEEE Access*, *8*, 133982-133994. https://doi.org/10.1109/ACCESS.2020.3010715

Ghorbani, S., Mahdaviani, K., Thaler, A., Kording, K., Cook, D. J., Blohm, G., & Troje, N. F. (2021). Movi: A Large Multi-Purpose Human Motion and Video Dataset. *PLOS ONE*, *16*(6), e0253157. https://doi.org/10.1371/journal.pone.0253157

Gianfrancesco, M. A., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, *178*(11), 1544-1547. https://doi.org/10.1001/jamainternmed.2018.3763

Ginis, P., Nieuwboer, A., Dorfman, M., Ferrari, A., Gazit, E., Canning, C. G., . . . Mirelman, A. (2016). Feasibility and Effects of Home-Based Smartphone-Delivered Automated Feedback Training for Gait in People with Parkinson's Disease: A Pilot Randomized Controlled Trial. *Parkinsonism & Related Disorders*, *22*, 28-34. https://doi.org/10.1016/j.parkreldis.2015.11.004

Gustafson, J. A., Robinson, M. E., Fitzgerald, G. K., Tashman, S., & Farrokhi, S. (2015). Knee Motion Variability in Patients with Knee Osteoarthritis: The Effect of Self-Reported Instability. *Clinical Biomechanics*, *30*(5), 475-480. https://doi.org/10.1016/j.clinbiomech.2015.03.007

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). Grade: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations. *Bmj*, *336*(7650), 924-926. https://doi.org/10.1136/bmj.39489.470347.ad

Gyllensten, I. C., & Bonomi, A. G. (2011). Identifying Types of Physical Activity with a Single Accelerometer: Evaluating Laboratory-Trained Algorithms in Daily Life. *IEEE transactions on biomedical engineering*, *58*(9), 2656-2663. https://doi.org/10.1109/tbme.2011.2160723

Hafer, J. F., Provenzano, S. G., Kern, K. L., Agresta, C. E., Grant, J. A., & Zernicke, R. F. (2020). Measuring Markers of Aging and Knee Osteoarthritis Gait Using Inertial Measurement Units. *Journal of Biomechanics*, *99*, 109567. https://doi.org/10.1016/j.jbiomech.2019.109567

Hall, M., Bennell, K. L., Wrigley, T. V., Metcalf, B. R., Campbell, P. K., Kasza, J., . . . Hinman, R. S. (2017). The Knee Adduction Moment and Knee Osteoarthritis Symptoms: Relationships According to Radiographic Disease Severity. *Osteoarthritis and Cartilage*, *25*(1), 34-41. https://doi.org/10.1016/j.joca.2016.08.014

Hall, M., Castelein, B., Wittoek, R., Calders, P., & Van Ginckel, A. (2019). Diet-Induced Weight Loss Alone or Combined with Exercise in Overweight or Obese People with Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Seminars in Arthritis and Rheumatism*, *48*(5), 765-777. https://doi.org/10.1016/j.semarthrit.2018.06.005

Hall, M., Hinman, R. S., Wrigley, T. V., Kasza, J., Lim, B. W., & Bennell, K. L. (2018). Knee Extensor Strength Gains Mediate Symptom Improvement in Knee Osteoarthritis: Secondary Analysis of a Randomised Controlled Trial. *Osteoarthritis and Cartilage*, *26*(4), 495-500. https://doi.org/10.1016/j.joca.2018.01.018

Hansen, M. M., Miron-Shatz, T., Lau, A. Y. S., & Paton, C. (2014). Big Data in Science and Healthcare: A Review of Recent Literature and Perspectives. *Yearb Med Inform*, *23*(01), 21-26. http://dx.doi.org/10.15265/IY-2014-0004

Harris-Hayes, M., Cornbleet, S. L., & Holtzman, G. W. (2010). Movement Impairment Syndromes of the Knee. In S. Sahrmann (Ed.), *Movement System Impairment Syndromes of the Extremities, Cervical and Thoracic Spines* (e-book ed., pp. 354-438). Elsevier Health Sciences.

Hart, H. F., Collins, N. J., Ackland, D. C., Cowan, S. M., & Crossley, K. M. (2015). Gait Characteristics of People with Lateral Knee Osteoarthritis after Acl Reconstruction. *Medicine & Science in Sports & Exercise*, *47*(11). https://doi.org/10.1249/MSS.0000000000000671

Hart, H. F., Collins, N. J., Ackland, D. C., & Crossley, K. M. (2015). Is Impaired Knee Confidence Related to Worse Kinesiophobia, Symptoms, and Physical Function in People with Knee Osteoarthritis after Anterior Cruciate Ligament Reconstruction? *Journal of Science and Medicine in Sport*, *18*(5), 512-517. https://doi.org/10.1016/j.jsams.2014.09.011

Hassan, N., Slight, R., Weiand, D., Vellinga, A., Morgan, G., Aboushareb, F., & Slight, S. P. (2021). Preventing Sepsis; How Can Artificial Intelligence Inform the Clinical Decision-Making Process? A Systematic Review. *International Journal of Medical Informatics*, *150*, 104457. https://doi.org/10.1016/j.ijmedinf.2021.104457

Hatfield, G. L., Stanish, W. D., & Hubley-Kozey, C. L. (2015). Three-Dimensional Biomechanical Gait Characteristics at Baseline Are Associated with Progression to Total Knee Arthroplasty. *Arthritis Care & Research*, *67*(7), 1004-1014. https://doi.org/10.1002/acr.22564

Hawker, G. A., Stewart, L., French, M. R., Cibere, J., Jordan, J. M., March, L., . . . Gooberman-Hill, R. (2008). Understanding the Pain Experience in Hip and Knee Osteoarthritis – an Oarsi/Omeract Initiative. *Osteoarthritis and Cartilage*, *16*(4), 415-422. https://doi.org/10.1016/j.joca.2007.12.017

He, Z., Liu, T., & Yi, J. (2019). A Wearable Sensing and Training System: Towards Gait Rehabilitation for Elderly Patients with Knee Osteoarthritis. *IEEE Sensors Journal*, *19*(14), 5936-5945. https://doi.org/10.1109/JSEN.2019.2908417

Heiden, T. L., Lloyd, D. G., & Ackland, T. R. (2009). Knee Joint Kinematics, Kinetics and Muscle Co-Contraction in Knee Osteoarthritis Patient Gait. *Clinical Biomechanics*, *24*(10), 833-841. https://doi.org/10.1016/j.clinbiomech.2009.08.005

Hekler, E. B., Klasnja, P., Chevance, G., Golaszewski, N. M., Lewis, D., & Sim, I. (2019). Why We Need a Small Data Paradigm. *BMC Medicine*, *17*(1), 133-133. https://doi,org/10.1186/s12916-019-1366-x

Hendry, D., Chai, K., Campbell, A., Hopper, L., O., O. S. P., & Straker, L. (2020). Development of a Human Activity Recognition System for Ballet Tasks. *Sports Medicine*, *6*(10). https://doi.org/10.1186/s40798-020-0237-5

Hendry, D., Napier, K., Hosking, R., Chai, K., Davey, P., Hopper, L., . . . Campbell, A. (2021). Development of a Machine Learning Model for the Estimation of Hip and Lumbar Angles in Ballet Dancers. *Medical Problems of Performing Artists*, *36*(2), 61-71. https://doi.org/10.21091/mppa.2021.2009

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A. Z., Muzny, M., Hartvigsen, G., Hopstock, L. A., & Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of Medical Internet Research*, *20*(3), e110. https://doi.org/10.2196/jmir.9157

Henriksen, M., Aaboe, J., & Bliddal, H. (2012). The Relationship between Pain and Dynamic Knee Joint Loading in Knee Osteoarthritis Varies with Radiographic Disease Severity. A Cross Sectional Study. *The Knee*, *19*(4), 392-398. https://doi.org/10.1016/j.knee.2011.07.003

Hensor, E. M. A., Dube, B., Kingsbury, S. R., Tennant, A., & Conaghan, P. G. (2015). Toward a Clinical Definition of Early Osteoarthritis: Onset of Patient-Reported Knee Pain Begins on Stairs. Data from the Osteoarthritis Initiative. *Arthritis Care & Research*, *67*(1), 40-47. https://doi.org/10.1002/acr.22418

Hernandez, V., Dadkhah, D., Babakeshizadeh, V., & Kulić, D. (2021). Lower Body Kinematics Estimation from Wearable Sensors for Walking and Running: A Deep Learning Approach. *Gait & Posture*, *83*, 185-193. https://doi.org/10.1016/j.gaitpost.2020.10.026

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li T, Page, M., & VA, W. (2020). Chapter 6: Choosing Effect Measures and Computing Estimates of Effect. In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (Updated September 2020)* (version 6.0 ed.). Cochrane. https://doi.org/ www.training.cochrane.org/handbook

Hillel, I., Gazit, E., Nieuwboer, A., Avanzino, L., Rochester, L., Cereatti, A., . . . Hausdorff, J. M. (2019). Is Every-Day Walking in Older Adults More Analogous to Dual-Task Walking or to Usual Walking? Elucidating the Gaps between Gait Performance in the Lab and During 24/7 Monitoring. *European Review of Aging and Physical Activity*, *16*(1), 6. https://doi.org/10.1186/s11556-019-0214-5

Hinman, R. S., Bennell, K. L., Metcalf, B. R., & Crossley, K. M. (2002). Delayed Onset of Quadriceps Activity and Altered Knee Joint Kinematics During Stair Stepping in Individuals with Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, *83*(8), 1080-1086. https://doi.org/10.1053/apmr.2002.33068

Hodges, P. W., & Smeets, R. J. (2015). Interaction between Pain, Movement, and Physical Activity: Short-Term Benefits, Long-Term Consequences, and Targets for Treatment. *The Clinical Journal of Pain*, *31*(2). https://doi.org/10.1097/AJP.0000000000000098

Hodges, P. W., & Tucker, K. (2011). Moving Differently in Pain: A New Theory to Explain the Adaptation to Pain. *PAIN*, *152*(3). https://doi.org/10.1016/j.pain.2010.10.020

Hodges, P. W., van den Hoorn, W., Wrigley, T. V., Hinman, R. S., Bowles, K.-A., Cicuttini, F., . . . Bennell, K. (2016). Increased Duration of Co-Contraction of Medial Knee Muscles Is Associated with Greater Progression of Knee Osteoarthritis. *Manual Therapy*, *21*, 151-158. https://doi.org/10.1016/j.math.2015.07.004

Holden, M. A., Button, K., Collins, N. J., Henrotin, Y., Hinman, R. S., Larsen, J. B., . . . Bennell, K. (2021). Guidance for Implementing Best Practice Therapeutic Exercise for Patients with Knee and Hip Osteoarthritis: What Does the Current Evidence Base Tell Us? *Arthritis Care & Research*, *73*(12), 1746-1753. https://doi.org/10.1002/acr.24434

Holsgaard-Larsen, A., Clausen, B., Søndergaard, J., Christensen, R., Andriacchi, T. P., & Roos, E. M. (2017). The Effect of Instruction in Analgesic Use Compared with Neuromuscular Exercise on Knee-Joint Load in Patients with Knee Osteoarthritis: A Randomized, Single-Blind, Controlled Trial. *Osteoarthritis and Cartilage*, *25*(4), 470-480. https://doi.org/10.1016/j.joca.2016.10.022

Horn, K. K., Jennings, S., Richardson, G., van Vliet, D., Hefford, C., & Abbott, J. H. (2012). The Patient-Specific Functional Scale: Psychometrics, Clinimetrics, and Application as a Clinical Outcome Measure. *Journal of Orthopaedic & Sports Physical Therapy*, *42*(1), 30-42. https://doi.org/10.2519/jospt.2012.3727

Hou, C. (2020). A Study on Imu-Based Human Activity Recognition Using Deep Learning and Traditional Machine Learning. 2020 5th International Conference on Computer and Communication Systems (ICCCS),

Huang, P.-C., Liu, K.-C., Hsieh, C.-Y., & Chan, C.-T. (2017). *Human Motion Identification for Rehabilitation Exercise Assessment of Knee Osteoarthritis* International Conference on Applied System Innovation (ICASI), Conference, 246-249,Location. https://doi.org/10.1109/ICASI.2017.7988396

Hunt, M. A., Charlton, J. M., Krowchuk, N. M., Tse, C. T. F., & Hatfield, G. L. (2018). Clinical and Biomechanical Changes Following a 4-Month Toe-out Gait Modification Program for People with Medial Knee Osteoarthritis: A Randomized Controlled Trial. *Osteoarthritis and Cartilage*, *26*(7), 903-911. https://doi.org/10.1016/j.joca.2018.04.010

Hunt, M. A., & Takacs, J. (2014). Effects of a 10-Week Toe-out Gait Modification Intervention in People with Medial Knee Osteoarthritis: A Pilot, Feasibility Study. *Osteoarthritis and Cartilage*, *22*(7), 904-911. https://doi.org/10.1016/j.joca.2014.04.007

Hunt, M. A., Takacs, J., Hart, K., Massong, E., Fuchko, K., & Biegler, J. (2014). Comparison of Mirror, Raw Video, and Real-Time Visual Biofeedback for Training Toe-out Gait in Individuals with Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, *95*(10), 1912-1917. https://doi.org/10.1016/j.apmr.2014.05.016

Hunt, M. A., Wrigley, T. V., Hinman, R. S., & Bennell, K. L. (2010). Individuals with Severe Knee Osteoarthritis (Oa) Exhibit Altered Proximal Walking Mechanics Compared with Individuals with Less Severe Oa and Those without Knee

Pain. *Arthritis Care & Research*, *62*(10), 1426-1432. https://doi.org/10.1002/acr.20248
Hunter, D. J. (2018). Osteoarthritis: Time for Us All to Shift the Needle. *Rheumatology*, *57*(suppl_4), iv1-iv2. 10.1093/rheumatology/key065
Hurley, M. V., Mitchell, H. L., & Walsh, N. (2003). In Osteoarthritis, the Psychosocial Benefits of Exercise Are as Important as Physiological Improvements. *Exercise and Sport Sciences Reviews*, *31*(3). https://doi.org/10.1097/00003677-200307000-00007
Hurley, M. V., Walsh, N. E., Mitchell, H. L., Pimm, T. J., Patel, A., Williamson, E., . . . Reeves, B. C. (2007). Clinical Effectiveness of a Rehabilitation Program Integrating Exercise, Self-Management, and Active Coping Strategies for Chronic Knee Pain: A Cluster Randomized Trial. *Arthritis Care & Research*, *57*(7), 1211-1219. https://doi.org/10.1002/art.22995
Hutting, N., Caneiro, J. P., Ong'wen, O. M., Miciak, M., & Roberts, L. (2022). Patient-Centered Care in Musculoskeletal Practice: Key Elements to Support Clinicians to Focus on the Person. *Musculoskeletal Science and Practice*, *57*, 102434. https://doi.org/10.1016/j.msksp.2021.102434
Iijima, H., Shimoura, K., Aoyama, T., & Takahashi, M. (2018). Biomechanical Characteristics of Stair Ambulation in Patients with Knee Oa: A Systematic Review with Meta-Analysis toward a Better Definition of Clinical Hallmarks. *Gait & Posture*, *62*, 191-201. https://doi.org/10.1016/j.gaitpost.2018.03.002
Jamshidi, A., Pelletier, J.-P., & Martel-Pelletier, J. (2019). Machine-Learning-Based Patient-Specific Prediction Models for Knee Osteoarthritis. *Nature Reviews Rheumatology*, *15*(1), 49-60. https://doi.org/10.1038/s41584-018-0130-5
Janidarmian, M., Roshan Fekr, A., Radecka, K., & Zilic, Z. (2017). A Comprehensive Analysis on Wearable Acceleration Sensors in Human Activity Recognition. *Sensors*, *17*(3), 529. https://doi.org/10.3390/s17030529
Jasper, L., Beaupre, L. A., Spence, J. C., & Jones, C. A. (2021). Validity of Tools to Measure Physical Activity in Older Adults Following Total Knee Arthroplasty. *Journal of Aging and Physical Activity*, *29*(4), 651-658. https://doi.org/10.1123/japa.2020-0240
Jenkyn, T. R., Hunt, M. A., Jones, I. C., Giffin, J. R., & Birmingham, T. B. (2008). Toe-out Gait in Patients with Knee Osteoarthritis Partially Transforms External Knee Adduction Moment into Flexion Moment During Early Stance Phase of Gait: A Tri-Planar Kinetic Mechanism. *Journal of Biomechanics*, *41*(2), 276-283. https://doi.org/10.1016/j.jbiomech.2007.09.015
Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., . . . Wang, Y. (2017). Artificial Intelligence in Healthcare: Past, Present and Future. *Stroke and Vascular Neurology*, *2*(4), 230-243. https://doi.org/10.1136/svn-2017-000101
Jiang, W., & Yin, Z. (2015). Human Activity Recognition Using Wearable Sensors by Deep Convolutional Neural Networks. Proceedings of the 23rd ACM international conference on Multimedia,
Jiang, X., Napier, C., Hannigan, B., Eng, J. J., & Menon, C. (2020). Estimating Vertical Ground Reaction Force During Walking Using a Single Inertial Sensor. *Sensors*, *20*(15), 4345. https://www.mdpi.com/1424-8220/20/15/4345
Jones, C. J., Rikli, R. E., & Beam, W. C. (1999). A 30-S Chair-Stand Test as a Measure of Lower Body Strength in Community-Residing Older Adults. *Research*

*Quarterly for Exercise and Sport*, *70*(2), 113-119. https://doi.org/10.1080/02701367.1999.10608028

Karatsidis, A., Jung, M., Schepers, H. M., Bellusci, G., de Zee, M., Veltink, P. H., & Andersen, M. S. (2019). Musculoskeletal Model-Based Inverse Dynamic Analysis under Ambulatory Conditions Using Inertial Motion Capture. *Medical Engineering & Physics*, *65*, 68-77. https://doi.org/10.1016/j.medengphy.2018.12.021

Karsdal, M. A., Christiansen, C., Ladel, C., Henriksen, K., Kraus, V. B., & Bay-Jensen, A. C. (2014). Osteoarthritis – a Case for Personalized Health Care? *Osteoarthritis and Cartilage*, *22*(1), 7-16. https://doi.org/10.1016/j.joca.2013.10.018

Kautz, T., Groh, B. H., Hannink, J., Jensen, U., Strubberg, H., & Eskofier, B. M. (2017). Activity Recognition in Beach Volleyball Using a Deep Convolutional Neural Network. *Data Mining and Knowledge Discovery*, *31*(6), 1678-1705. https://doi.org/10.1007/s10618-017-0495-0

Kellgren, J. H., & Lawrence, J. S. (1957). Radiological Assessment of Osteo-Arthrosis. *Annals of the Rheumatic Diseases*, *16*(4), 494-502. https://doi.org/10.1136/ard.16.4.494

Kent, P., Laird, R., & Haines, T. (2015). The Effect of Changing Movement and Posture Using Motion-Sensor Biofeedback, Versus Guidelines-Based Care, on the Clinical Outcomes of People with Sub-Acute or Chronic Low Back Pain-a Multicentre, Cluster-Randomised, Placebo-Controlled, Pilot Trial. *BMC Musculoskeletal Disorders*, *16*(1), 131. https://doi.org/10.1186/s12891-015-0591-5

Kent, P., Sullivan, P., Smith, A., Haines, T., Campbell, A., McGregor, A. H., . . . Hancock, M. (2019). Restore—Cognitive Functional Therapy with or without Movement Sensor Biofeedback Versus Usual Care for Chronic, Disabling Low Back Pain: Study Protocol for a Randomised Controlled Trial. *BMJ Open*, *9*(8), e031133. https://doi.org/10.1136/bmjopen-2019-031133

Kim, E., Helal, S., & Cook, D. (2010). Human Activity Recognition and Pattern Discovery. *IEEE Pervasive Computing*, *9*(1), 48-53. https://doi.org/10.1109/MPRV.2010.7

King, L. K., Birmingham, T. B., Kean, C. O., Jones, I. C., Bryant, D. M., & Giffin, J. R. (2008). Resistance Training for Medial Compartment Knee Osteoarthritis and Malalignment. *Medicine & Science in Sports & Exercise*, *40*(8), 1376-1384. https://doi.org/10.1249/MSS.0b013e31816f1c4a

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *arXiv*, *arXiv:1412.6980v9*. https://doi.org/10.48550/arXiv.1412.6980

Kiss, R. M. (2011). Effect of Severity of Knee Osteoarthritis on the Variability of Gait Parameters. *Journal of Electromyography and Kinesiology*, *21*(5), 695-703. https://doi.org/10.1016/j.jelekin.2011.07.011

Kitchin, R., & McArdle, G. (2016). What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets. *Big Data & Society*, *3*(1), 2053951716631130. https://doi.org/10.1177/2053951716631130

Kittelson, A. J., George, S. Z., Maluf, K. S., & Stevens-Lapsley, J. E. (2014). Future Directions in Painful Knee Osteoarthritis: Harnessing Complexity in a

Heterogeneous Population. *Physical Therapy*, *94*(3), 422-432. https://doi.org/10.2522/ptj.20130256

Knoop, J., van der Leeden, M., Thorstensson, C. A., Roorda, L. D., Lems, W. F., Knol, D. L., . . . Dekker, J. (2011). Identification of Phenotypes with Different Clinical Outcomes in Knee Osteoarthritis: Data from the Osteoarthritis Initiative. *Arthritis Care & Research*, *63*(11), 1535-1542. https://doi.org/10.1002/acr.20571

Kobsar, D., & Ferber, R. (2018). Wearable Sensor Data to Track Subject-Specific Movement Patterns Related to Clinical Outcomes Using a Machine Learning Approach. *Sensors*, *18*(9), 2828. https://doi.org/10.3390/s18092828

Kobsar, D., Masood, Z., Khan, H., Khalil, N., Kiwan, M. Y., Ridd, S., & Tobis, M. (2020). Wearable Inertial Sensors for Gait Analysis in Adults with Osteoarthritis—a Scoping Review. *Sensors*, *20*(24), 7143. https://doi.org/10.3390/s20247143

Kobsar, D., Osis, S. T., Boyd, J. E., Hettinga, B. A., & Ferber, R. (2017). Wearable Sensors to Predict Improvement Following an Exercise Intervention in Patients with Knee Osteoarthritis. *Journal of NeuroEngineering and Rehabilitation*, *14*(1), 94. https://doi.org/10.1186/s12984-017-0309-z

Kok, M., Hol, J. D., & Schön, T. B. (2017). Using Inertial Sensors for Position and Orientation Estimation. *Foundations and Trends® in Signal Processing*, *11*(1-2), 1-153. https://doi.org/10.1561/2000000094

Kokkotis, C., Moustakidis, S., Papageorgiou, E., Giakas, G., & Tsaopoulos, D. E. (2020). Machine Learning in Knee Osteoarthritis: A Review. *Osteoarthritis and Cartilage Open*, *2*(3), 100069. https://doi.org/10.1016/j.ocarto.2020.100069

Kongsted, A., Kent, P., Quicke, J. G., Skou, S. T., & Hill, J. C. (2020). Risk-Stratified and Stepped Models of Care for Back Pain and Osteoarthritis: Are We Heading Towards a Common Model? *Pain reports*, *5*(5), e843-e843. https://doi.org/10.1097/PR9.0000000000000843

Koning, B. H. W., van der Krogt, M. M., Baten, C. T. M., & Koopman, B. F. J. M. (2015). Driving a Musculoskeletal Model with Inertial and Magnetic Measurement Units. *Computer Methods in Biomechanics and Biomedical Engineering*, *18*(9), 1003-1013. https://doi.org/10.1080/10255842.2013.867481

Konrath, J. M., Karatsidis, A., Schepers, H. M., Bellusci, G., de Zee, M., & Andersen, M. S. (2019). Estimation of the Knee Adduction Moment and Joint Contact Force During Daily Living Activities Using Inertial Motion Capture. *Sensors*, *19*(7). https://doi.org/10.3390/s19071681

Kowalski, K., Rhodes, R., Naylor, P.-J., Tuokko, H., & MacDonald, S. (2012). Direct and Indirect Measurement of Physical Activity in Older Adults: A Systematic Review of the Literature. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 148. https://doi.org/10.1186/1479-5868-9-148

Kraus, V. B., Blanco, F. J., Englund, M., Karsdal, M. A., & Lohmander, L. S. (2015). Call for Standardized Definitions of Osteoarthritis and Risk Stratification for Clinical Trials and Clinical Use. *Osteoarthritis and Cartilage*, *23*(8), 1233-1241. https://doi.org/10.1016/j.joca.2015.03.036

Krause, D. A., Boyd, M. S., Hager, A. N., Smoyer, E. C., Thompson, A. T., & Hollman, J. H. (2015). Reliability and Accuracy of a Goniometer Mobile Device Application for Video Measurement of the Functional Movement

Screen Deep Squat Test. *International Journal of Sports Physical Therapy*, *10*(1), 37-44. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4325286/

Laird, R. A., Keating, J. L., Ussing, K., Li, P., & Kent, P. (2019). Does Movement Matter in People with Back Pain? Investigating 'Atypical' Lumbo-Pelvic Kinematics in People with and without Back Pain Using Wireless Movement Sensors. *BMC Musculoskeletal Disorders*, *20*(1), 28. https://doi/org/10.1186/s12891-018-2387-x

Laird, R. A., Kent, P., & Keating, J. L. (2012). Modifying Patterns of Movement in People with Low Back Pain -Does It Help? A Systematic Review. *BMC Musculoskeletal Disorders*, *13*(1), 169. https://doi.org/10.1186/1471-2474-13-169

Lane, N. E., Brandt, K., Hawker, G., Peeva, E., Schreyer, E., Tsuji, W., & Hochberg, M. C. (2011). Oarsi-Fda Initiative: Defining the Disease State of Osteoarthritis. *Osteoarthritis and Cartilage*, *19*(5), 478-482. https://doi.org/10.1016/j.joca.2010.09.013

Laney, D. (2001). 3d Data Management: Controlling Data Volume, Velocity and Variety. *META group research note*, *6*(70), 1.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lee, C. J., & Lee, J. K. (2022). Inertial Motion Capture-Based Wearable Systems for Estimation of Joint Kinetics: A Systematic Review. *Sensors*, *22*(7), 2507. https://www.mdpi.com/1424-8220/22/7/2507

Lee, J., Joo, H., Lee, J., & Chee, Y. (2020). Automatic Classification of Squat Posture Using Inertial Sensors: Deep Learning Approach. *Sensors*, *20*(2), 361. https://doi.org/10.3390/s20020361

Lehman, G. J. (2018). The Role and Value of Symptom-Modification Approaches in Musculoskeletal Practice. *Journal of Orthopaedic & Sports Physical Therapy*, *48*(6), 430-435. https://doi.org/10.2519/jospt.2018.0608

Lentz, T. A., George, S. Z., Manickas-Hill, O., Malay, M. R., O'Donnell, J., Jayakumar, P., . . . Mather III, R. C. (2020). What General and Pain-Associated Psychological Distress Phenotypes Exist among Patients with Hip and Knee Osteoarthritis? *Clinical Orthopaedics and Related Research*, *478*(12), 2768. https://doi.org/10.1097%2FCORR.0000000000001520

Leslie, H. H., Zhou, X., Spiegelman, D., & Kruk, M. E. (2018). Health System Measurement: Harnessing Machine Learning to Advance Global Health. *PLOS ONE*, *13*(10), e0204958. https://doi.org/10.1371/journal.pone.0204958

Lim, B.-W., Hinman, R. S., Wrigley, T. V., Sharma, L., & Bennell, K. L. (2008). Does Knee Malalignment Mediate the Effects of Quadriceps Strengthening on Knee Adduction Moment, Pain, and Function in Medial Knee Osteoarthritis? A Randomized Controlled Trial. *Arthritis Care & Research*, *59*(7), 943-951. https://doi.org/10.1002/art.23823

Lim, H., Kim, B., & Park, S. (2020). Prediction of Lower Limb Kinetics and Kinematics During Walking by a Single Imu on the Lower Back Using Machine Learning. *Sensors*, *20*(1). https://doi.org/10.3390/s20010130

Lin, I., Wiles, L., Waller, R., Goucke, R., Nagree, Y., Gibberd, M., . . . O'Sullivan, P. P. B. (2020). What Does Best Practice Care for Musculoskeletal Pain Look Like? Eleven Consistent Recommendations from High-Quality Clinical

Practice Guidelines: Systematic Review. *British Journal of Sports Medicine*, *54*(2), 79. https://doi.org/10.1136/bjsports-2018-099878

Lin, S. (2022). A Clinician's Guide to Artificial Intelligence (Ai): Why and How Primary Care Should Lead the Health Care Ai Revolution. *The Journal of the American Board of Family Medicine*, *35*(1), 175. https://doi.org/10.3122/jabfm.2022.01.210226

Lipperts, M., van Laarhoven, S., Senden, R., Heyligers, I., & Grimm, B. (2017). Clinical Validation of a Body-Fixed 3d Accelerometer and Algorithm for Activity Monitoring in Orthopaedic Patients. *Journal of Orthopaedic Translation*, *11*, 19-29. https://doi.org/10.1016/j.jot.2017.02.003

Liu, H., & Lang, B. (2019). Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. *Applied Sciences*, *9*(20). https://doi.org/10.3390/app9204396

Lo, G. H., Harvey, W. F., & McAlindon, T. E. (2012). Associations of Varus Thrust and Alignment with Pain in Knee Osteoarthritis. *Arthritis & Rheumatism*, *64*(7), 2252-2259. https://doi.org/10.1002/art.3442

Lo, G. H., McAlindon, T. E., Niu, J., Zhang, Y., Beals, C., Dabrowski, C., . . . Hunter, D. J. (2009). Bone Marrow Lesions and Joint Effusion Are Strongly and Independently Associated with Weight-Bearing Pain in Knee Osteoarthritis: Data from the Osteoarthritis Initiative. *Osteoarthritis and Cartilage*, *17*(12), 1562-1569. https://doi.org/10.1016/j.joca.2009.06.006

Lonini, L., Gupta, A., Kording, K., & Jayaraman, A. (2016, 16-20 August). *Activity Recognition in Patients with Lower Limb Impairments: Do We Need Training Data from Each Patient?* 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Conference, 3265-3268,Location. https://doi.org/10.1109/EMBC.2016.7591425

Machado, G. P. M., Gignac, M. A. M., & Badley, E. M. (2008). Participation Restrictions among Older Adults with Osteoarthritis: A Mediated Model of Physical Symptoms, Activity Limitations, and Depression. *Arthritis Care & Research*, *59*(1), 129-135. https://doi.org/10.1002/art.23259

Mahmoudian, A., Van Assche, D., Herzog, W., & Luyten, F. P. (2018). Towards Secondary prevention Of early knee osteoarthritis. *RMD Open*, *4*(2), e000468. https://doi.org/10.1136/rmdopen-2017-000468

Maly, M. R. (2009). Linking Biomechanics to Mobility and Disability in People with Knee Osteoarthritis. *Exercise and Sport Sciences Reviews*, *37*(1). https://doi.org/10.1097/JES.0b013e3181912071

Maly, M. R., & Krupa, T. (2007). Personal Experience of Living with Knee Osteoarthritis among Older Adults. *Disability and Rehabilitation*, *29*(18), 1423-1433. https://doi.org/10.1080/09638280601029985

Marriott, K. A., Birmingham, T. B., Leitch, K. M., Pinto, R., & Giffin, J. R. (2019). Strong Independent Associations between Gait Biomechanics and Pain in Patients with Knee Osteoarthritis. *Journal of Biomechanics*, *94*, 123-129. https://doi.org/10.1016/j.jbiomech.2019.07.015

Martinez-Hernandez, U., & Dehghani-Sanij, A. A. (2018). Adaptive Bayesian Inference System for Recognition of Walking Activities and Prediction of Gait Events Using Wearable Sensors. *Neural Networks*, *102*, 107-119. https://doi.org/10.1016/j.neunet.2018.02.017

Martinez-Hernandez, U., & Dehghani-Sanij, A. A. (2019). Probabilistic Identification of Sit-to-Stand and Stand-to-Sit with a Wearable Sensor. *Pattern Recognition Letters*, *118*, 32-41. https://doi.org/10.1016/j.patrec.2018.03.020

Marvel, F. A., Wang, J., & Martin, S. S. (2018). Digital Health Innovation: A Toolkit to Navigate from Concept to Clinical Testing. *JMIR Cardio*, *2*(1), e2. https://doi.org/10.2196/cardio.7586

Mazumder, M., Banbury, C., Yao, X., Karlaš, B., Rojas, W. G., Diamos, S., . . . Jurado, D. (2022). Dataperf: Benchmarks for Data-Centric Ai Development. *arXiv preprint arXiv:2207.10062*. https://doi.org/10.48550/arXiv.2207.10062

McAlindon, T. E., Driban, J. B., Henrotin, Y., Hunter, D. J., Jiang, G. L., Skou, S. T., . . . Schnitzer, T. (2015). Oarsi Clinical Trials Recommendations: Design, Conduct, and Reporting of Clinical Trials for Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *23*(5), 747-760. https://doi.org/10.1016/j.joca.2015.03.005

McCarthy, I., Hodgins, D., Mor, A., Elbaz, A., & Segal, G. (2013). Analysis of Knee Flexion Characteristics and How They Alter with the Onset of Knee Osteoarthritis: A Case Control Study. *BMC Musculoskeletal Disorders*, *14*(1), 169. https://doi.org/10.1186/1471-2474-14-169

McGinley, J. L., Baker, R., Wolfe, R., & Morris, M. E. (2009). The Reliability of Three-Dimensional Kinematic Gait Measurements: A Systematic Review. *Gait & Posture*, *29*(3), 360-369. https://doi.org/10.1016/j.gaitpost.2008.09.003

McGrath, J., Neville, J., Stewart, T., & Cronin, J. (2020). Upper Body Activity Classification Using an Inertial Measurement Unit in Court and Field-Based Sports: A Systematic Review. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, *235*(2), 83-95. https://doi.org/10.1177/1754337120959754

McQuade, K. J., & de Oliveira, A. S. (2011). Effects of Progressive Resistance Strength Training on Knee Biomechanics During Single Leg Step-up in Persons with Mild Knee Osteoarthritis. *Clinical Biomechanics*, *26*(7), 741-748. https://doi.org/10.1016/j.clinbiomech.2011.03.006

Messier, S. P., Loeser, R. F., Hoover, J. L., Semble, E. L., & Wise, C. M. (1992). Osteoarthritis of the Knee: Effects on Gait, Strength, and Flexibility. *Archives of Physical Medicine and Rehabilitation*, *73*(1), 29-36. https://doi.org/10.5555/uri:pii:0003999392902221

Middel, B., & van Sonderen, E. (2002). Statistical Significant Change Versus Relevant or Important Change in (Quasi) Experimental Design: Some Conceptual and Methodological Problems in Estimating Magnitude of Intervention-Related Change in Health Services Research. *International journal of integrated care*, *2*, e15-e15. https://doi.org/10.5334/ijic.65

Mihalko, S. L., Cox, P., Beavers, D. P., Miller, G. D., Nicklas, B. J., Lyles, M., . . . Messier, S. P. (2019). Effect of Intensive Diet and Exercise on Self-Efficacy in Overweight and Obese Adults with Knee Osteoarthritis: The Idea Randomized Clinical Trial. *Translational Behavioral Medicine*, *9*(2), 227-235. https://doi.org/10.1093/tbm/iby037

Milanese, S., Gordon, S., Buettner, P., Flavell, C., Ruston, S., Coe, D., . . . McCormack, S. (2014). Reliability and Concurrent Validity of Knee Angle Measurement: Smart Phone App Versus Universal Goniometer Used by

Experienced and Novice Clinicians. *Manual Therapy*, *19*(6), 569-574. https://doi.org/10.1016/j.math.2014.05.009

Mills, K., Hunt, M. A., Leigh, R., & Ferber, R. (2013). A Systematic Review and Meta-Analysis of Lower Limb Neuromuscular Alterations Associated with Knee Osteoarthritis During Level Walking. *Clinical Biomechanics*, *28*(7), 713-724. https://doi.org/10.1016/j.clinbiomech.2013.07.008

Mitchell, T. M. (1997). *Machine Learning*. New York : McGraw-Hill.

Miyazaki, T., Wada, M., Kawahara, H., Sato, M., Baba, H., & Shimada, S. (2002). Dynamic Load at Baseline Can Predict Radiographic Disease Progression in Medial Compartment Knee Osteoarthritis. *Annals of the Rheumatic Diseases*, *61*(7), 617. http://dx.doi.org/10.1136/ard.61.7.617

Mobasheri, A., & Batt, M. (2016). An Update on the Pathophysiology of Osteoarthritis. *Annals of Physical and Rehabilitation Medicine*, *59*(5), 333-339. https://doi.org/10.1016/j.rehab.2016.07.004

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The Prisma Statement. *Ann Intern Med*, *151*(4), 264-269, w264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Montoye, H. J., Washburn, R., Servais, S., Ertl, A., Webster, J. G., & Nagle, F. J. (1983). Estimation of Energy Expenditure by a Portable Accelerometer. *Medicine & Science in Sports & Exercise*, *15*(5), 403-407. https://pubmed.ncbi.nlm.nih.gov/6645869/

Moola, S., Munn, Z., Tufanaru, C., Aromataris, E., Sears, K., Sfetcu, R., . . . Lisy, K. (2017). Chapter 7: Systematic Reviews of Etiology and Risk. In *Joanna Briggs Institute Reviewer's Manual*. The Joanna Briggs Institute. https://doi.org/https://doi.org/10.46658/JBIMES-20-08

Moyer, R. F., Birmingham, T. B., Bryant, D. M., Giffin, J. R., Marriott, K. A., & Leitch, K. M. (2015). Biomechanical Effects of Valgus Knee Bracing: A Systematic Review and Meta-Analysis. *Osteoarthritis and Cartilage*, *23*(2), 178-188. https://doi.org/10.1016/j.joca.2014.11.018

Mundt, M., Johnson, W. R., Potthast, W., Markert, B., Mian, A., & Alderson, J. (2021). A Comparison of Three Neural Network Approaches for Estimating Joint Angles and Moments from Inertial Measurement Units. *Sensors*, *21*(13), 4535. https://doi.org/10.3390/s21134535

Mundt, M., Koeppe, A., Bamer, F., David, S., & Markert, B. (2020). Artificial Neural Networks in Motion Analysis—Applications of Unsupervised and Heuristic Feature Selection Techniques. *Sensors*, *20*(16), 4581. https://doi.org/10.3390/s20164581

Mundt, M., Koeppe, A., David, S., Bamer, F., Potthast, W., & Markert, B. (2020). Prediction of Ground Reaction Force and Joint Moments Based on Optical Motion Capture Data During Gait. *Medical Engineering & Physics*, *86*, 29-34. https://doi.org/10.1016/j.medengphy.2020.10.001

Mundt, M., Koeppe, A., David, S., Witter, T., Bamer, F., Potthast, W., & Markert, B. (2020). Estimation of Gait Mechanics Based on Simulated and Measured Imu Data Using an Artificial Neural Network. *Frontiers in Bioengineering and Biotechnology*, *8*(41). https://doi.org/10.3389/fbioe.2020.00041

Mundt, M., Thomsen, W., Witter, T., Koeppe, A., David, S., Bamer, F., . . . Markert, B. (2020). Prediction of Lower Limb Joint Angles and Moments During Gait Using Artificial Neural Networks. *Medical & Biological Engineering & Computing*, *58*(1), 211-225. https://doi.org/10.1007/s11517-019-02061-3

National Centre for Biotechnology Information. (2021). *Medical Subject Headings 2021* https://meshb.nlm.nih.gov/record/ui?ui=D005081

National Institute for Health & Care Excellence. (2014). *Osteoarthritis: Care and Management* (CG177) https://www.nice.org.uk/guidance/ng226

National Library of Medicine. (2019). *Welcome to Medical Subject Headings*. https://www.nlm.nih.gov/mesh/meshhome.html

Nebel, M. B., Sims, E. L., Keefe, F. J., Kraus, V. B., Guilak, F., Caldwell, D. S., . . . Schmitt, D. (2009). The Relationship of Self-Reported Pain and Functional Impairment to Gait Mechanics in Overweight and Obese Persons with Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, *90*(11), 1874-1879. https://doi.org/10.1016/j.apmr.2009.07.010

Ng, A. (2021). *A Chat with Andrew on Mlops: From Model-Centric to Data-Centric Ai*. https://youtu.be/06-AZXmwHjo

Ngueleu, A. M., Blanchette, A. K., Maltais, D., Moffet, H., McFadyen, B. J., Bouyer, L., & Batcho, C. S. (2019). Validity of Instrumented Insoles for Step Counting, Posture and Activity Recognition: A Systematic Review. *Sensors*, *19*(11). https://doi.org/10.3390/s19112438

Nguyen, H., Lebel, K., Bogard, S., Goubault, E., Boissy, P., & Duval, C. (2017). Using Inertial Sensors to Automatically Detect and Segment Activities of Daily Living in People with Parkinson's Disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(1), 197-204. https://doi.org/10.1109/TNSRE.2017.2745418

Nie, Y., Wang, H., Xu, B., Zhou, Z., Shen, B., & Pei, F. (2019). The Relationship between Knee Adduction Moment and Knee Osteoarthritis Symptoms According to Static Alignment and Pelvic Drop. *BioMed Research International*, *2019*, 7603249. https://doi.org/10.1155/2019/7603249

Nodehi Moghadam, A., Rahnama, L., Noorizadeh Dehkordi, S., & Abdollahi, S. (2020). Exercise Therapy May Affect Scapular Position and Motion in Individuals with Scapular Dyskinesis: A Systematic Review of Clinical Trials. *Journal of Shoulder and Elbow Surgery*, *29*(1), e29-e36. https://doi.org/10.1016/j.jse.2019.05.037

Noorbergen, T. J., Adam, M. T., Roxburgh, M., & Teubner, T. (2021). Co-Design in Mhealth Systems Development: Insights from a Systematic Literature Review. *AIS Transactions on Human-Computer Interaction*, *13*(2), 175-205. https://doi.org/10.17705/1thci.00147

O'Connell, M., Farrokhi, S., & Fitzgerald, G. K. (2016). The Role of Knee Joint Moments and Knee Impairments on Self-Reported Knee Pain During Gait in Patients with Knee Osteoarthritis. *Clinical Biomechanics*, *31*, 40-46. https://doi.org/10.1016/j.clinbiomech.2015.10.003

O'Reilly, M., Caulfield, B., Ward, T., Johnston, W., & Doherty, C. (2018). Wearable Inertial Sensor Systems for Lower Limb Exercise Detection and Evaluation: A Systematic Review. *Sports Medicine*, *48*(5), 1221-1246. https://doi.org/10.1007/s40279-018-0878-4

O'Sullivan, P. B., Caneiro, J. P., O'Keeffe, M., Smith, A., Dankaerts, W., Fersum, K., & O'Sullivan, K. (2018). Cognitive Functional Therapy: An Integrated Behavioral Approach for the Targeted Management of Disabling Low Back Pain. *Physical Therapy*, *98*(5), 408-423. https://doi.org/10.1093/ptj/pzy022

Øiestad, B. E., Juhl, C. B., Culvenor, A. G., Berg, B., & Thorlund, J. B. (2022). Knee Extensor Muscle Weakness Is a Risk Factor for the Development of Knee Osteoarthritis: An Updated Systematic Review and Meta-Analysis Including 46 819 Men and Women. *British Journal of Sports Medicine*, *56*(6), 349. https://doi.org/10.1136/bjsports-2021-104861

Oubre, B., Lane, S., Holmes, S., Boyer, K., & Lee, S. I. (2021). Estimating Ground Reaction Force and Center of Pressure Using Low-Cost Wearable Devices. *IEEE transactions on biomedical engineering*, *69*(4), 1461-1468. https://doi.org/10.1109/TBME.2021.3120346

Oxford English Dictionary. *"Artificial Intelligence, N."*. Oxford University Press. https://www.oed.com/view/Entry/271625?redirectedFrom=artificial+intelligence

Papi, E., Belsi, A., & McGregor, A. H. (2015). A Knee Monitoring Device and the Preferences of Patients Living with Osteoarthritis: A Qualitative Study. *BMJ Open*, *5*(9), e007980. https://doi.org/10.1136/bmjopen-2015-007980

Papi, E., Murtagh, G. M., & McGregor, A. H. (2016). Wearable Technologies in Osteoarthritis: A Qualitative Study of Clinicians' Preferences. *BMJ Open*, *6*(1), e009544. https://doi.org/10.1136/bmjopen-2015-009544

Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing Bias in Artificial Intelligence in Health Care. *JAMA*, *322*(24), 2377-2378. https://doi.org/10.1001/jama.2019.18058

Patel, M. S., Asch, D. A., & Volpp, K. G. (2015). Wearable Devices as Facilitators, Not Drivers, of Health Behavior Change. *JAMA*, *313*(5), 459-460. https://doi.org/10.1001/jama.2014.14781

Pfister, A., West, A. M., Bronner, S., & Noah, J. A. (2014). Comparative Abilities of Microsoft Kinect and Vicon 3d Motion Capture for Gait Analysis. *Journal of Medical Engineering & Technology*, *38*(5), 274-280. https://doi.org/10.3109/03091902.2014.909540

Picerno, P. (2017). 25 Years of Lower Limb Joint Kinematics by Using Inertial and Magnetic Sensors: A Review of Methodological Approaches. *Gait & Posture*, *51*, 239-246. https://doi.org/10.1016/j.gaitpost.2016.11.008

Poitras, I., Dupuis, F., Bielmann, M., Campeau-Lecours, A., Mercier, C., Bouyer, L. J., & Roy, J.-S. (2019). Validity and Reliability of Wearable Sensors for Joint Angle Estimation: A Systematic Review. *Sensors*, *19*(7). https://doi.org/10.3390/s19071555

Poulsen, E., Goncalves, G. H., Bricca, A., Roos, E. M., Thorlund, J. B., & Juhl, C. B. (2019). Knee Osteoarthritis Risk Is Increased 4-6 Fold after Knee Injury – a Systematic Review and Meta-Analysis. *British Journal of Sports Medicine*, *53*(23), 1454. https://doi.org/10.1136/bjsports-2018-100022

Preece, S. J., Brookes, N., Williams, A. E., Jones, R. K., Starbuck, C., Jones, A., & Walsh, N. E. (2021). A New Integrated Behavioural Intervention for Knee Osteoarthritis: Development and Pilot Study. *BMC Musculoskeletal Disorders*, *22*(1), 526. https://doi.org/10.1186/s12891-021-04389-0

Preece, S. J., Jones, R. K., Brown, C. A., Cacciatore, T. W., & Jones, A. K. P. (2016). Reductions in Co-Contraction Following Neuromuscular Re-Education in People with Knee Osteoarthritis. *BMC Musculoskeletal Disorders*, *17*(1), 372. https://doi.org/10.1186/s12891-016-1209-2

Previtali, D., Andriolo, L., Di Laura Frattura, G., Boffa, A., Candrian, C., Zaffagnini, S., & Filardo, G. (2020). Pain Trajectories in Knee Osteoarthritis—a Systematic Review and Best Evidence Synthesis on Pain Predictors. *Journal of Clinical Medicine*, *9*(9). https://doi.org/10.3390/jcm9092828

Qi, J., Yang, P., Waraich, A., Deng, Z., Zhao, Y., & Yang, Y. (2018). Examining Sensor-Based Physical Activity Recognition and Monitoring for Healthcare Using Internet of Things: A Systematic Review. *Journal of Biomedical Informatics*, *87*, 138-153. https://doi.org/10.1016/j.jbi.2018.09.002

Qiu, S., Zhao, H., Jiang, N., Wang, Z., Liu, L., An, Y., . . . Fortino, G. (2022). Multi-Sensor Information Fusion Based on Machine Learning for Real Applications in Human Activity Recognition: State-of-the-Art and Research Challenges. *Information Fusion*, *80*, 241-265. https://doi.org/10.1016/j.inffus.2021.11.006

Radzimski, A. O., Mündermann, A., & Sole, G. (2012). Effect of Footwear on the External Knee Adduction Moment — a Systematic Review. *The Knee*, *19*(3), 163-175. https://doi.org/10.1016/j.knee.2011.05.013

Ramanujam, E., Perumal, T., & Padmavathi, S. (2021). Human Activity Recognition with Smartphone and Wearable Sensors Using Deep Learning Techniques: A Review. *IEEE Sensors Journal*, *21*(12), 13029-13040. https://doi.org/10.1109/JSEN.2021.3069927

Rapp, E., Shin, S., Thomsen, W., Ferber, R., & Halilaj, E. (2021). Estimation of Kinematics from Inertial Measurement Units Using a Combined Deep Learning and Optimization Framework. *Journal of Biomechanics*, *116*, 110229. https://doi.org/10.1016/j.jbiomech.2021.110229

Rast, F. M., & Labruyère, R. (2020). Systematic Review on the Application of Wearable Inertial Sensors to Quantify Everyday Life Motor Activity in People with Mobility Impairments. *Journal of NeuroEngineering and Rehabilitation*, *17*(1), 148. https://doi.org/10.1186/s12984-020-00779-y

Rausch Osthoff, A. K., Niedermann, K., Braun, J., Adams, J., Brodin, N., Dagfinrud, H., . . . Vliet Vlieland, T. P. M. (2018). 2018 Eular Recommendations for Physical Activity in People with Inflammatory Arthritis and Osteoarthritis. *Ann Rheum Dis*, *77*(9), 1251-1260. https://doi.org/10.1136/annrheumdis-2018-213585

Ren, L., Jones, R. K., & Howard, D. (2008). Whole Body Inverse Dynamics over a Complete Gait Cycle Based Only on Measured Kinematics. *Journal of Biomechanics*, *41*(12), 2750-2759. https://doi.org/10.1016/j.jbiomech.2008.06.001

Renani, M. S., Eustace, A. M., Myers, C. A., & Clary, C. W. (2021). The Use of Synthetic Imu Signals in the Training of Deep Learning Models Significantly Improves the Accuracy of Joint Kinematic Predictions. *Sensors*, *21*(17), 5876. https://doi.org/10.3390/s21175876

Renani, M. S., Myers, C. A., Zandie, R., Mahoor, M. H., Davidson, B. S., & Clary, C. W. (2020). Deep Learning in Gait Parameter Prediction for Oa and Tka

Patients Wearing Imu Sensors. *Sensors*, *20*(19), 5553. https://doi.org/10.3390/s20195553

Renggli, D., Graf, C., Tachatos, N., Singh, N., Meboldt, M., Taylor, W. R., . . . Schmid Daners, M. (2020). Wearable Inertial Measurement Units for Assessing Gait in Real-World Environments. *Frontiers in Physiology*, *11*(90). https://doi.org/10.3389/fphys.2020.00090

Richards, J. G. (1999). The Measurement of Human Motion: A Comparison of Commercially Available Systems. *Human Movement Science*, *18*(5), 589-602. https://doi.org/10.1016/S0167-9457(99)00023-8

Richards, R., van den Noort, J. C., Dekker, J., & Harlaar, J. (2017). Gait Retraining with Real-Time Biofeedback to Reduce Knee Adduction Moment: Systematic Review of Effects and Methods Used. *Archives of Physical Medicine and Rehabilitation*, *98*(1), 137-150. https://doi.org/10.1016/j.apmr.2016.07.006

Richards, R., van den Noort, J. C., van der Esch, M., Booij, M. J., & Harlaar, J. (2018). Gait Retraining Using Real-Time Feedback in Patients with Medial Knee Osteoarthritis: Feasibility and Effects of a Six-Week Gait Training Program. *The Knee*, *25*(5), 814-824. https://doi.org/10.1016/j.knee.2018.05.014

Robert-Lachaine, X., Mecheri, H., Larue, C., & Plamondon, A. (2017). Validation of Inertial Measurement Units with an Optoelectronic System for Whole-Body Motion Analysis. *Medical & Biological Engineering & Computing*, *55*(4), 609-619. 10.1007/s11517-016-1537-2

Robert-Lachaine, X., Mecheri, H., Muller, A., Larue, C., & Plamondon, A. (2020). Validation of a Low-Cost Inertial Motion Capture System for Whole-Body Motion Analysis. *Journal of Biomechanics*, *99*, 109520. https://doi.org/10.1016/j.jbiomech.2019.109520

Robles-García, V., Corral-Bergantiños, Y., Espinosa, N., Jácome, M. A., García-Sancho, C., Cudeiro, J., & Arias, P. (2015). Spatiotemporal Gait Patterns During Overt and Covert Evaluation in Patients with Parkinson's Disease and Healthy Subjects: Is There a Hawthorne Effect? *Journal of Applied Biomechanics*, *31*(3), 189. https://doi.org/10.1123/jab.2013-0319

Rodríguez-Martín, D., Samà, A., Pérez-López, C., Català, A., Moreno Arostegui, J. M., Cabestany, J., . . . Rodríguez-Molinero, A. (2017). Home Detection of Freezing of Gait Using Support Vector Machines through a Single Waist-Worn Triaxial Accelerometer. *PLOS ONE*, *12*(2), e0171764. https://doi.org/10.1371/journal.pone.0171764

Roman-Blas, J. A., Mendoza-Torres, L. A., Largo, R., & Herrero-Beaumont, G. (2020). Setting up Distinctive Outcome Measures for Each Osteoarthritis Phenotype. *Therapeutic Advances in Musculoskeletal Disease*, *12*, 1759720X20937966. https://doi.org/10.1177/1759720X20937966

Roos, E. M., & Lohmander, L. S. (2003). The Knee Injury and Osteoarthritis Outcome Score (Koos): From Joint Injury to Osteoarthritis. *Health and Quality of Life Outcomes*, *1*(1), 64. https://doi.org/10.1186/1477-7525-1-64

Roper, J. A., Bressel, E., & Tillman, M. D. (2013). Acute Aquatic Treadmill Exercise Improves Gait and Pain in People with Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, *94*(3), 419-425. https://doi.org/10.1016/j.apmr.2012.10.027

Royal Australian College of General Practitioners. (2018). *Guideline for the Management of Knee and Hip Osteoarthritis*. https://www.racgp.org.au/clinical-resources/clinical-guidelines/key-racgp-guidelines/view-all-racgp-guidelines/knee-and-hip-osteoarthritis

Runhaar, J., Luijsterburg, P., Dekker, J., & Bierma-Zeinstra, S. M. A. (2015). Identifying Potential Working Mechanisms Behind the Positive Effects of Exercise Therapy on Pain and Function in Osteoarthritis; a Systematic Review. *Osteoarthritis and Cartilage*, *23*(7), 1071-1082. https://doi.org/10.1016/j.joca.2014.12.027

Rutherford, D., Baker, M., Wong, I., & Stanish, W. (2017). The Effect of Age and Knee Osteoarthritis on Muscle Activation Patterns and Knee Joint Biomechanics During Dual Belt Treadmill Gait. *Journal of Electromyography and Kinesiology*, *34*, 58-64. https://doi.org/10.1016/j.jelekin.2017.04.001

Rutherford, D. J., Hubley-Kozey, C. L., & Stanish, W. D. (2012). Knee Effusion Affects Knee Mechanics and Muscle Activity During Gait in Individuals with Knee Osteoarthritis. *Osteoarthritis and Cartilage*, *20*(9), 974-981. https://doi.org/10.1016/j.joca.2012.05.014

Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., & Kording, K. P. (2017). The Need to Approximate the Use-Case in Clinical Machine Learning. *GigaScience*, *6*(5), gix019. https://doi.org/10.1093/gigascience/gix019

Safiri, S., Kolahi, A.-A., Smith, E., Hill, C., Bettampadi, D., Mansournia, M. A., . . . Cross, M. (2020). Global, Regional and National Burden of Osteoarthritis 1990-2017: A Systematic Analysis of the Global Burden of Disease Study 2017. *Annals of the Rheumatic Diseases*, *79*(6), 819. https://doi.org/10.1136/annrheumdis-2019-216515

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210-229. https://doi.org/10.1147/rd.33.0210

Sani, S., Massie, S., Wiratunga, N., & Cooper, K. (2017). Learning Deep and Shallow Features for Human Activity Recognition. Knowledge Science, Engineering and Management, Cham.

Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, *2*(6), 420. https://doi.org/10.1007/s42979-021-00815-1

Schache, A. G., Fregly, B. J., Crossley, K. M., Hinman, R. S., & Pandy, M. G. (2008). The Effect of Gait Modification on the External Knee Adduction Moment Is Reference Frame Dependent. *Clinical Biomechanics*, *23*(5), 601-608. https://doi.org/10.1016/j.clinbiomech.2007.12.008

Schall, M. C., Fethke, N. B., Chen, H., Oyama, S., & Douphrate, D. I. (2016). Accuracy and Repeatability of an Inertial Measurement Unit System for Field-Based Occupational Studies. *Ergonomics*, *59*(4), 591-602. https://doi.org/10.1080/00140139.2015.1079335

Segal, N. A., Boyer, E. R., Wallace, R., Torner, J. C., & Yack, H. J. (2013). Association between Chair Stand Strategy and Mobility Limitations in Older Adults with Symptomatic Knee Osteoarthritis. *Archives of Physical Medicine and Rehabilitation*, *94*(2), 375-383. https://doi.org/10.1016/j.apmr.2012.09.026

Segal, N. A. M. D. M. S., Glass, N. A. M. A., Teran-Yengle, P. P., Singh, B. P., Wallace, R. B. M. D., & Yack, H. J. P. (2015). Intensive Gait Training for Older Adults with Symptomatic Knee Osteoarthritis. *American Journal of Physical Medicine & Rehabilitation*, *94(10 Suppl 1)*(Issue), 848-858. https://doi.org/10.1097/PHM.0000000000000264

Senders, J. T., Staples, P. C., Karhade, A. V., Zaki, M. M., Gormley, W. B., Broekman, M. L. D., . . . Arnaout, O. (2018). Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurgery*, *109*, 476-486.e471. https://doi.org/10.1016/j.wneu.2017.09.149

Seror, R., Tubach, F., Baron, G., Falissard, B., Logeart, I., Dougados, M., & Ravaud, P. (2008). Individualising the Western Ontario and Mcmaster Universities Osteoarthritis Index (Womac) Function Subscale: Incorporating Patient Priorities for Improvement to Measure Functional Impairment in Hip or Knee Osteoarthritis. *Annals of the Rheumatic Diseases*, *67*(4), 494. http://dx.doi.org/10.1136/ard.2007.074591

Sharif, B., Garner, R., Sanmartin, C., Flanagan, W. M., Hennessy, D., & Marshall, D. A. (2016). Risk of Work Loss Due to Illness or Disability in Patients with Osteoarthritis: A Population-Based Cohort Study. *Rheumatology*, *55*(5), 861-868. https://doi.org/10.1093/rheumatology/kev428

Shen, C.-L., James, C. R., Chyu, M.-C., Bixby, W. R., Jean-Michel, B., Zumwalt, M. A., & Poklikuha, G. (2008). Effects of Tai Chi on Gait Kinematics, Physical Function, and Pain in Elderly with Knee Osteoarthritis — a Pilot Study. *American Journal of Chinese Medicine*, *36*(02), 219-232. https://doi.org/10.1142/s0192415x08005734

Shull, P. B., Shultz, R., Silder, A., Dragoo, J. L., Besier, T. F., Cutkosky, M. R., & Delp, S. L. (2013). Toe-in Gait Reduces the First Peak Knee Adduction Moment in Patients with Medial Compartment Knee Osteoarthritis. *Journal of Biomechanics*, *46*(1), 122-128. https://doi.org/10.1016/j.jbiomech.2012.10.019

Shull, P. B., Silder, A., Shultz, R., Dragoo, J. L., Besier, T. F., Delp, S. L., & Cutkosky, M. R. (2013). Six-Week Gait Retraining Program Reduces Knee Adduction Moment, Reduces Pain, and Improves Function for Individuals with Medial Compartment Knee Osteoarthritis. *Journal of Orthopaedic Research*, *31*(7), 1020-1025. https://doi.org/10.1002/jor.22340

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The Performance of Lstm and Bilstm in Forecasting Time Series. 2019 IEEE International Conference on Big Data (Big Data),

Sidey-Gibbons, J. A. M., & Sidey-Gibbons, C. J. (2019). Machine Learning in Medicine: A Practical Introduction. *BMC Medical Research Methodology*, *19*(1), 64. https://doi.org/10.1186/s12874-019-0681-4

Silverwood, V., Blagojevic-Bucknall, M., Jinks, C., Jordan, J. L., Protheroe, J., & Jordan, K. P. (2015). Current Evidence on Risk Factors for Knee Osteoarthritis in Older Adults: A Systematic Review and Meta-Analysis. *Osteoarthritis and Cartilage*, *23*(4), 507-515. https://doi.org/10.1016/j.joca.2014.11.019

Simic, M., Wrigley, T. V., Hinman, R. S., Hunt, M. A., & Bennell, K. L. (2013). Altering Foot Progression Angle in People with Medial Knee Osteoarthritis: The Effects of Varying Toe-in and Toe-out Angles Are Mediated by

Pain and Malalignment. *Osteoarthritis and Cartilage*, *21*(9), 1272-1280. https://doi.org/10.1016/j.joca.2013.06.001

Sims, E. L., Keefe, F. J., Kraus, V. B., Guilak, F., Queen, R. M., & Schmitt, D. (2009). Racial Differences in Gait Mechanics Associated with Knee Osteoarthritis. *Aging Clinical and Experimental Research*, *21*(6), 463-469. https://doi.org/10.1007/BF03327442

Skou, S., Pedersen, B. K., Abbott, J. H., Patterson, B., & Barton, C. (2018). Physical Activity and Exercise Therapy Benefit More Than Just Symptoms and Impairments in People with Hip and Knee Osteoarthritis. *Journal of Orthopaedic and Sports Physical Therapy*, *48*(6), 439-447. https://doi.org/10.2519/jospt.2018.7877

Skou, S. T., Rasmussen, S., Laursen, M. B., Rathleff, M. S., Arendt-Nielsen, L., Simonsen, O., & Roos, E. M. (2015). The Efficacy of 12 Weeks Non-Surgical Treatment for Patients Not Eligible for Total Knee Replacement: A Randomized Controlled Trial with 1-Year Follow-Up. *Osteoarthritis and Cartilage*, *23*(9), 1465-1475. https://doi.org/10.1016/j.joca.2015.04.021

Skou, S. T., Rasmussen, S., Simonsen, O., & Roos, E. M. (2015). Knee Confidence as It Relates to Self-Reported and Objective Correlates of Knee Osteoarthritis: A Cross-Sectional Study of 220 Patients. *Journal of Orthopaedic & Sports Physical Therapy*, *45*(10), 765-771. https://doi.org/10.2519/jospt.2015.5864

Skou, S. T., & Roos, E. M. (2017). Good Life with Osteoarthritis in Denmark (Gla:D™): Evidence-Based Education and Supervised Neuromuscular Exercise Delivered by Certified Physiotherapists Nationwide. *BMC Musculoskeletal Disorders*, *18*(1), 72. https://doi.org/10.1186/s12891-017-1439-y

Skou, S. T., Wrigley, T. V., Metcalf, B. R., Hinman, R. S., & Bennell, K. L. (2014). Association of Knee Confidence with Pain, Knee Instability, Muscle Strength, and Dynamic Varus–Valgus Joint Motion in Knee Osteoarthritis. *Arthritis Care & Research*, *66*(5), 695-701. https://doi.org/10.1002/acr.22208

Sled, E. A., Khoja, L., Deluzio, K. J., Olney, S. J., & Culham, E. G. (2010). Effect of a Home Program of Hip Abductor Exercises on Knee Joint Loading, Strength, Function, and Pain in People with Knee Osteoarthritis: A Clinical Trial. *Physical Therapy*, *90*(6), 895-904. https://doi.org/10.2522/ptj.20090294

Small, S. R., Bullock, G. S., Khalid, S., Barker, K., Trivella, M., & Price, A. J. (2019). Current Clinical Utilisation of Wearable Motion Sensors for the Assessment of Outcome Following Knee Arthroplasty: A Scoping Review. *BMJ Open*, *9*(12), e033832. https://doi.org/10.1136/bmjopen-2019-033832

Somvanshi, M., Chavan, P., Tambade, S., & Shinde, S. (2016). *A Review of Machine Learning Techniques Using Decision Tree and Support Vector Machine* International conference on computing communication control and automation (ICCUBEA), Conference, 1-7,Location. https://doi.org/10.1109/ICCUBEA.2016.7860040

Song-Mi, L., Sang Min, Y., & Heeryon, C. (2017, 13-16 February). *Human Activity Recognition from Accelerometer Data Using Convolutional Neural Network* IEEE International Conference on Big Data and Smart Computing (BigComp), Conference, 131-134,Location. https://doi.org/10.1109/BIGCOMP.2017.7881728

Sonoo, M., Iijima, H., & Kanemura, N. (2019). Altered Sagittal Plane Kinematics and Kinetics During Sit-to-Stand in Individuals with Knee Osteoarthritis: A Systematic Review and Meta-Analysis. *Journal of Biomechanics*, *96*, 109331. https://doi.org/10.1016/j.jbiomech.2019.109331

Sparkes, V., Whatling, G. M., Biggs, P., Khatib, N., Al-Amri, M., Williams, D., . . . Holt, C. (2019). Comparison of Gait, Functional Activities, and Patient-Reported Outcome Measures in Patients with Knee Osteoarthritis and Healthy Adults Using 3d Motion Analysis and Activity Monitoring: An Exploratory Case-Control Analysis. *Orthopedic Research and Reviews*, *11*, 129-140. http://doi.org/10.2147/ORR.S199107

Stetter, B. J., Krafft, F. C., Ringhof, S., Stein, T., & Sell, S. (2020). A Machine Learning and Wearable Sensor Based Approach to Estimate External Knee Flexion and Adduction Moments During Various Locomotion Tasks. *Frontiers in Bioengineering and Biotechnology*, *8*(9). https://doi.org/10.3389/fbioe.2020.00009

Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2019). Estimation of Knee Joint Forces in Sport Movements Using Wearable Sensors and Machine Learning. *Sensors*, *19*(17). https://doi.org/10.3390/s19173690

Stevens-Lapsley, J. E., Schenkman, M. L., & Dayton, M. R. (2011). Comparison of Self-Reported Knee Injury and Osteoarthritis Outcome Score to Performance Measures in Patients after Total Knee Arthroplasty. *PM&R*, *3*(6), 541-549. https://doi.org/10.1016/j.pmrj.2011.03.002

Stratford, P. W., Kennedy, D., Pagura, S. M., & Gollish, J. D. (2003). The Relationship between Self-Report and Performance-Related Measures: Questioning the Content Validity of Timed Tests. *Arthritis Care & Research*, *49*(4), 535-540. https://doi.org/10.1002/art.11196

Stratford, P. W., & Kennedy, D. M. (2006). Performance Measures Were Necessary to Obtain a Complete Picture of Osteoarthritic Patients. *Journal of Clinical Epidemiology*, *59*(2), 160-167. https://doi.org/10.1016/j.jclinepi.2005.07.012

Sung, J., Han, S., Park, H., Cho, H.-M., Hwang, S., Park, J. W., & Youn, I. (2022). Prediction of Lower Extremity Multi-Joint Angles During Overground Walking by Using a Single Imu with a Low Frequency Based on an Lstm Recurrent Neural Network. *Sensors*, *22*(1), 53. https://doi.org/10.3390/s22010053

Swain, S., Sarmanova, A., Mallen, C., Kuo, C. F., Coupland, C., Doherty, M., & Zhang, W. (2020). Trends in Incidence and Prevalence of Osteoarthritis in the United Kingdom: Findings from the Clinical Practice Research Datalink (Cprd). *Osteoarthritis and Cartilage*, *28*(6), 792-801. https://doi.org/10.1016/j.joca.2020.03.004

Tan, J.-S., Beheshti, B. K., Binnie, T., Davey, P., Caneiro, J. P., Kent, P., . . . Campbell, A. (2021). Human Activity Recognition for People with Knee Osteoarthritis—a Proof-of-Concept. *Sensors*, *21*(10). https://doi.org/10.3390/s21103381

Tan, J.-S., Tikoft, E., O'Sullivan, P., Smith, A., Campbell, A., Caneiro, J. P., & Kent, P. (2021). The Relationship between Changes in Movement and Activity Limitation or Pain in People with Knee Osteoarthritis: A Systematic Review.

*Journal of Orthopaedic & Sports Physical Therapy*, *51*(10), 492-502. https://doi.org/10.2519/jospt.2021.10418

Tan, J.-S., Tippaya, S., Binnie, T., Davey, P., Napier, K., Caneiro, J. P., . . . Campbell, A. (2022). Predicting Knee Joint Kinematics from Wearable Sensor Data in People with Knee Osteoarthritis and Clinical Considerations for Future Machine Learning Models. *Sensors*, *22*(2). https://doi.org/10.3390/s22020446

Teo, P. L., Bennell, K. L., Lawford, B. J., Egerton, T., Dziedzic, K. S., & Hinman, R. S. (2020). Physiotherapists May Improve Management of Knee Osteoarthritis through Greater Psychosocial Focus, Being Proactive with Advice, and Offering Longer-Term Reviews: A Qualitative Study. *Journal of Physiotherapy*, *66*(4), 256-265. https://doi.org/10.1016/j.jphys.2020.09.005

Terwee, C. B., Mokkink, L. B., Steultjens, M. P. M., & Dekker, J. (2006). Performance-Based Methods for Measuring the Physical Function of Patients with Osteoarthritis of the Hip or Knee: A Systematic Review of Measurement Properties. *Rheumatology*, *45*(7), 890-902. https://doi.org/10.1093/rheumatology/kei267

Terwee, C. B., van der Slikke, R. M. A., van Lummel, R. C., Benink, R. J., Meijers, W. G. H., & de Vet, H. C. W. (2006). Self-Reported Physical Functioning Was More Influenced by Pain Than Performance-Based Physical Functioning in Knee-Osteoarthritis Patients. *Journal of Clinical Epidemiology*, *59*(7), 724-731. https://doi.org/10.1016/j.jclinepi.2005.11.019

Teufl, W., Miezal, M., Taetz, B., Fröhlich, M., & Bleser, G. (2019). Validity of Inertial Sensor Based 3d Joint Kinematics of Static and Dynamic Sport and Physiotherapy Specific Movements. *PLOS ONE*, *14*(2), e0213064. https://doi.org/10.1371/journal.pone.0213064

Thorp, L., Wimmer, M., Foucher, K., Sumner, D., Shakoor, N., & Block, J. (2010). The Biomechanical Effects of Focused Muscle Training on Medial Knee Loads in Oa of the Knee: A Pilot, Proof of Concept Study. *Journal of Musculoskeletal and Neuronal Interactions*, *10*(2), 166-173. http://www.ismni.org/jmni/pdf/40/07THORP.pdf

Thorp, L. E., Sumner, D. R., Block, J. A., Moisio, K. C., Shott, S., & Wimmer, M. A. (2006). Knee Joint Loading Differs in Individuals with Mild Compared with Moderate Medial Knee Osteoarthritis. *Arthritis & Rheumatism*, *54*(12), 3842-3849. https://doi.org/10.1002/art.22247

Thorstensson, C. A., Henriksson, M., von Porat, A., Sjödahl, C., & Roos, E. M. (2007). The Effect of Eight Weeks of Exercise on Knee Adduction Moment in Early Knee Osteoarthritis – a Pilot Study. *Osteoarthritis and Cartilage*, *15*(10), 1163-1170. https://doi.org/10.1016/j.joca.2007.03.012

Tolk, J. J., Janssen, R. P. A., Prinsen, C. A. C., Latijnhouwers, D. A. J. M., van der Steen, M. C., Bierma-Zeinstra, S. M. A., & Reijman, M. (2019). The Oarsi Core Set of Performance-Based Measures for Knee Osteoarthritis Is Reliable but Not Valid and Responsive. *Knee Surgery, Sports Traumatology, Arthroscopy*, *27*(9), 2898-2909. https://doi.org/10.1007/s00167-017-4789-y

Turcot, K., Aissaoui, R., Boivin, K., Pelletier, M., Hagemeister, N., & de Guise, J. A. (2009). The Responsiveness of Three-Dimensional Knee Accelerations Used as an Estimation of Knee Instability and Loading Transmission During Gait in

Osteoarthritis Patient's Follow-Up. *Osteoarthritis and Cartilage*, *17*(2), 213-219. https://doi.org/10.1016/j.joca.2008.06.021

Turcot, K., Armand, S., Fritschy, D., Hoffmeyer, P., & Suvà, D. (2012). Sit-to-Stand Alterations in Advanced Knee Osteoarthritis. *Gait & Posture*, *36*(1), 68-72. https://doi.org/10.1016/j.gaitpost.2012.01.005

Turing, A. M. (1950). I.—Computing Machinery and Intelligence. *Mind*, *LIX*(236), 433-460. https://doi.org/10.1093/mind/LIX.236.433

Um, T. T., Pfister, F. M. J., Pichler, D., Endo, S., Lang, M., Hirche, S., . . . Kulić, D. (2017). *Data Augmentation of Wearable Sensor Data for Parkinson's Disease Monitoring Using Convolutional Neural Networks* Proceedings of the 19th ACM International Conference on Multimodal Interaction, Conference, 216–220,Location. https://doi.org/10.1145/3136755.3136817

Valkenet, K., & Veenhof, C. (2019). Validity of Three Accelerometers to Investigate Lying, Sitting, Standing and Walking. *PLOS ONE*, *14*(5), e0217545. https://doi.org/10.1371/journal.pone.0217545

van den Noort, J., van der Esch, M., Steultjens, M. P. M., Dekker, J., Schepers, M., Veltink, P. H., & Harlaar, J. (2013). Ambulatory Measurement of the Knee Adduction Moment in Patients with Osteoarthritis of the Knee. *Journal of Biomechanics*, *46*(1), 43-49. https://doi.org/10.1016/j.jbiomech.2012.09.030

van der Esch, M., & Dekker, J. (2014). Neuromuscular Mechanisms Explaining Functional Decline. In *Exercise and Physical Functioning in Osteoarthritis* (pp. 49-67). Springer.

van der Straaten, R., De Baets, L., Jonkers, I., & Timmermans, A. (2018). Mobile Assessment of the Lower Limb Kinematics in Healthy Persons and in Persons with Degenerative Knee Disorders: A Systematic Review. *Gait & Posture*, *59*, 229-241. https://doi.org/10.1016/j.gaitpost.2017.10.005

van der Straaten, R., Wesseling, M., Jonkers, I., Vanwanseele, B., Bruijnes, A. K. B. D., Malcorps, J., . . . Timmermans, A. (2020). Functional Movement Assessment by Means of Inertial Sensor Technology to Discriminate between Movement Behaviour of Healthy Controls and Persons with Knee Osteoarthritis. *Journal of NeuroEngineering and Rehabilitation*, *17*(1), 65. https://doi.org/10.1186/s12984-020-00694-2

van Dijk, G. M., Dekker, J., Veenhof, C., & van den Ende, C. H. M. (2006). Course of Functional Status and Pain in Osteoarthritis of the Hip or Knee: A Systematic Review of the Literature. *Arthritis Care & Research*, *55*(5), 779-785. https://doi.org/10.1002/art.22244

van Spil, W. E., Bierma-Zeinstra, S. M. A., Deveza, L. A., Arden, N. K., Bay-Jensen, A. C., Kraus, V. B., . . . Hunter, D. J. (2020). A Consensus-Based Framework for Conducting and Reporting Osteoarthritis Phenotype Research. *Arthritis Research & Therapy*, *22*(1), 54. https://doi.org/10.1186/s13075-020-2143-0

van Tunen, J. A. C., Dell'Isola, A., Juhl, C., Dekker, J., Steultjens, M., Thorlund, J. B., & Lund, H. (2018). Association of Malalignment, Muscular Dysfunction, Proprioception, Laxity and Abnormal Joint Loading with Tibiofemoral Knee Osteoarthritis - a Systematic Review and Meta-Analysis. *BMC Musculoskeletal Disorders*, *19*(1), 273. https://doi.org/10.1186/s12891-018-2202-8

Verlaan, L., Boekesteijn, R. J., Oomen, P. W., Liu, W.-Y., Peters, M. J. M., Witlox, M. A., . . . Meijer, K. (2018). Biomechanical Alterations During Sit-to-Stand

Transfer Are Caused by a Synergy between Knee Osteoarthritis and Obesity. *BioMed Research International*, *2018*, 3519498. https://doi.org/10.1155/2018/3519498

Verlaan, L., Bolink, S. A. A. N., Van Laarhoven, S. N., Lipperts, M., Heyligers, I. C., Grimm, B., & Senden, R. (2015). Accelerometer-Based Physical Activity Monitoring in Patients with Knee Osteoarthritis: Objective and Ambulatory Assessment of Actual Physical Activity During Daily Life Circumstances. *The open biomedical engineering journal*, *9*, 157-163. https://doi.org/10.2174/1874120701509010157

Vicon. (2021). *What Is Motion Capture*. Retrieved June 1 from https://www.vicon.com/about-us/what-is-motion-capture/

Vlaeyen, J. W. S., & Linton, S. J. (2000). Fear-Avoidance and Its Consequences in Chronic Musculoskeletal Pain: A State of the Art. *PAIN*, *85*(3), 317-332. https://doi.org/10.1016/S0304-3959(99)00242-0

Vlaeyen, J. W. S., & Linton, S. J. (2012). Fear-Avoidance Model of Chronic Musculoskeletal Pain: 12 Years On. *PAIN*, *153*(6). https://doi.org/10.1016/j.pain.2011.12.009

Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., . . . Aboyans, V. (2012). Years Lived with Disability (Ylds) for 1160 Sequelae of 289 Diseases and Injuries 1990–2010: A Systematic Analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2163-2196. https://doi.org/10.1016/S0140-6736(12)61729-2

Wallis, J. A., Taylor, N. F., Bunzli, S., & Shields, N. (2019). Experience of Living with Knee Osteoarthritis: A Systematic Review of Qualitative Studies. *BMJ Open*, *9*(9), e030060. https://doi.org/10.1136/bmjopen-2019-030060

Wang, C., Chan, P. P. K., Lam, B. M. F., Wang, S., Zhang, J. H., Chan, Z. Y. S., . . . Cheung, R. T. H. (2020). Real-Time Estimation of Knee Adduction Moment for Gait Retraining in Patients with Knee Osteoarthritis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(4), 888-894. https://doi.org/10.1109/TNSRE.2020.2978537

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep Learning for Sensor-Based Activity Recognition: A Survey. *Pattern Recognition Letters*, *119*, 3-11. https://doi.org/10.1016/j.patrec.2018.02.010

Wang, J., Siddicky, S. F., Oliver, T. E., Dohm, M. P., Barnes, C. L., & Mannen, E. M. (2019). Biomechanical Changes Following Knee Arthroplasty During Sit-to-Stand Transfers: Systematic Review. *The Journal of Arthroplasty*, *34*(10), 2494-2501. https://doi./org/10.1016/j.arth.2019.05.028

Wang, P., Yang, L., Li, H., Lei, Z., Yang, X., Liu, C., . . . He, C. (2016). Effects of Whole-Body Vibration Training with Quadriceps Strengthening Exercise on Functioning and Gait Parameters in Patients with Medial Compartment Knee Osteoarthritis: A Randomised Controlled Preliminary Study. *Physiotherapy*, *102*(1), 86-92. https://doi.org/10.1016/j.physio.2015.03.3720

Wang, S., Chan, P. P. K., Lam, B. M. F., Chan, Z. Y. S., Zhang, J. H. W., Wang, C., . . . Cheung, R. T. H. (2021). Sensor-Based Gait Retraining Lowers Knee Adduction Moment and Improves Symptoms in Patients with Knee Osteoarthritis: A Randomized Controlled Trial. *Sensors*, *21*(16). https://doi.org/10.3390/s21165596

Weiss, A., Sharifi, S., Plotnik, M., van Vugt, J. P. P., Giladi, N., & Hausdorff, J. M. (2011). Toward Automated, at-Home Assessment of Mobility among Patients with Parkinson Disease, Using a Body-Worn Accelerometer. *Neurorehabilitation and Neural Repair*, *25*(9), 810-818. https://doi.org/10.1177/1545968311424869

Wernli, K., O'Sullivan, P., Smith, A., Campbell, A., & Kent, P. (2020). Movement, Posture and Low Back Pain. How Do They Relate? A Replicated Single-Case Design in 12 People with Persistent, Disabling Low Back Pain. *European Journal of Pain*, *24*(9), 1831-1849. https://doi.org/10.1002/ejp.1631

Wernli, K., Tan, J.-S., O'Sullivan, P., Smith, A., Campbell, A., & Kent, P. (2020). Does Movement Change When Low Back Pain Changes? A Systematic Review. *Journal of Orthopaedic & Sports Physical Therapy*, *50*(12), 664-670. https://doi.org/10.2519/jospt.2020.9635

Wernli, K., Tan, J.-S., O'Sullivan, P., Smith, A., Campbell, A., & Kent, P. (2021). The Relationship between Changes in Movement and Changes in Low Back Pain: A Systematic Review of Single-Case Designs. *JOSPT Cases*, *1*(4), 199-219. https://doi.org/10.2519/josptcases.2021.10231

Weygers, I., Kok, M., Konings, M., Hallez, H., De Vroey, H., & Claeys, K. (2020). Inertial Sensor-Based Lower Limb Joint Kinematics: A Methodological Systematic Review. *Sensors*, *20*(3). https://doi.org/10.3390/s20030673

Whatling, G. M., Evans, S. L., & Holt, C. A. (2010). Introducing a New Staircase Design to Quantify Healthy Knee Function During Stair Ascent and Descent. *Computer Methods in Biomechanics and Biomedical Engineering*, *13*(3), 371-378. https://doi.org/10.1080/10255840903251296

Whiteside, D., Cant, O., Connolly, M., & Reid, M. (2017). Monitoring Hitting Load in Tennis Using Inertial Sensors and Machine Learning. *International Journal of Sports Physiology and Performance*, *12*(9), 1212-1217. https://doi.org/10.1123/ijspp.2016-0683

Wiemken, T. L., & Kelley, R. R. (2020). Machine Learning in Epidemiology and Health Outcomes Research. *Annual Review of Public Health*, *41*(1), 21-36. https://doi.org/10.1146/annurev-publhealth-040119-094437

Wilkie, R., Peat, G., Thomas, E., & Croft, P. (2007). Factors Associated with Restricted Mobility Outside the Home in Community-Dwelling Adults Ages Fifty Years and Older with Knee Pain: An Example of Use of the International Classification of Functioning to Investigate Participation Restriction. *Arthritis Care & Research*, *57*(8), 1381-1389. https://doi.org/10.1002/art.23083

Wilson, J., Amiri, P., & Brandon, S. (2021). Tibiofemoral Joint Contact Forces During Gait Differ with Radiographic Knee Oa Progression over 3 Years. *Osteoarthritis and Cartilage*, *29*, S176-S178. https://doi.org/10.1016/j.joca.2021.02.244

World Health Organization. (2002). *Towards a Common Language for Functioning, Disability, and Health: Icf* (The international classification of functioning, disability and health, Issue. https://cdn.who.int/media/docs/default-source/classification/icf/icfbeginnersguide.pdf

Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., van Beijnum, B.-J. F., & Veltink, P. H. (2018). Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors. *Frontiers in Physiology*, *9*(218). https://doi.org/10.3389/fphys.2018.00218

Wright, A. A., Cook, C. E., Baxter, G. D., Dockerty, J. D., & Abbott, J. H. (2011). A Comparison of 3 Methodological Approaches to Defining Major Clinically Important Improvement of 4 Performance Measures in Patients with Hip Osteoarthritis. *Journal of Orthopaedic & Sports Physical Therapy*, *41*(5), 319-327. https://doi.org/10.2519/jospt.2011.3515

Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., . . . Stokes, I. (2002). Isb Recommendation on Definitions of Joint Coordinate System of Various Joints for the Reporting of Human Joint Motion—Part I: Ankle, Hip, and Spine. *Journal of Biomechanics*, *35*(4), 543-548. https://doi.org/10.1016/S0021-9290(01)00222-6

Xu, C., Chai, D., He, J., Zhang, X., & Duan, S. (2019). Innohar: A Deep Neural Network for Complex Human Activity Recognition. *IEEE Access*, *7*, 9893-9902. https://doi.org/10.1109/ACCESS.2018.2890675

Xu, W., Pang, Y., Yang, Y., & Liu, Y. (2018, 20-24 August). *Human Activity Recognition Based on Convolutional Neural Network* 24th International Conference on Pattern Recognition (ICPR), Conference, 165-170,Location. https://doi.org/10.1109/ICPR.2018.8545435

Zanardi, A. P. J., da Silva, E. S., Costa, R. R., Passos-Monteiro, E., dos Santos, I. O., Kruel, L. F. M., & Peyré-Tartaruga, L. A. (2021). Gait Parameters of Parkinson's Disease Compared with Healthy Controls: A Systematic Review and Meta-Analysis. *Scientific Reports*, *11*(1), 752. https://doi.org/10.1038/s41598-020-80768-2

Zeni, J. A., & Higginson, J. S. (2009). Differences in Gait Parameters between Healthy Subjects and Persons with Moderate and Severe Knee Osteoarthritis: A Result of Altered Walking Speed? *Clinical Biomechanics*, *24*(4), 372-378. https://doi.org/10.1016/j.clinbiomech.2009.02.001

Zhang, W., Doherty, M. A., Peat, G., Bierma-Zeinstra, M., Arden, N., Bresnihan, B., . . . Lohmander, L. (2010). Eular Evidence-Based Recommendations for the Diagnosis of Knee Osteoarthritis. *Annals of the Rheumatic Diseases*, *69*(3), 483-489. https://doi.org/10.1136/ard.2009.113100

Zheng, S., Tu, L., Cicuttini, F., Zhu, Z., Han, W., Antony, B., . . . Ding, C. (2021). Depression in Patients with Knee Osteoarthritis: Risk Factors and Associations with Joint Symptoms. *BMC Musculoskeletal Disorders*, *22*(1), 40. https://doi.org/10.1186/s12891-020-03875-1

Zhu, Q., Huang, L., Wu, X., Wang, L., Zhang, Y., Fang, M., . . . Li, J. X. (2016). Effects of Tai Ji Quan Training on Gait Kinematics in Older Chinese Women with Knee Osteoarthritis: A Randomized Controlled Trial. *Journal of Sport and Health Science*, *5*(3), 297-303. https://doi.org/10.1016/j.jshs.2016.02.003

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

# Thesis Appendices

## Ethics Approval

Curtin University

Office of Research and Development

GPO Box U1987
Perth Western Australia 6845

**Telephone** +61 8 9266 7863
**Facsimile** +61 8 9266 3793
**Web** research.curtin.edu.au

04-Oct-2017

Name: Amity Campbell
Department/School: School of Physiotherapy and Exercise Science
Email: A.Campbell@curtin.edu.au

Dear Amity Campbell

**RE: Ethics Office approval**
**Approval number: HRE2017-0695**

Thank you for submitting your application to the Human Research Ethics Office for the project **An exploration of the association between movement and pain-related thoughts, beliefs and emotions in knee osteoarthritis.**.

Your application was reviewed through the Curtin University Low risk review process.

The review outcome is: **Approved**.

Your proposal meets the requirements described in the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research (2007)*.

Approval is granted for a period of one year from **04-Oct-2017** to **03-Oct-2018**. Continuation of approval will be granted on an annual basis following submission of an annual report.

Personnel authorised to work on this project:

| Name | Role |
|---|---|
| Campbell, Amity | CI |
| Ng, Leo | Supervisor |
| Smith, Anne | Supervisor |
| O'Sullivan, Peter | Supervisor |
| Kent, Peter | Supervisor |
| West, Tara | Student |

Approved documents:

| Document |
|---|

**Standard conditions of approval**

1. Research must be conducted according to the approved proposal
2. Report in a timely manner anything that might warrant review of ethical approval of the project including:
   - proposed changes to the approved proposal or conduct of the study
   - unanticipated problems that might affect continued ethical acceptability of the project
   - major deviations from the approved proposal and/or regulatory guidelines
   - serious adverse events
3. Amendments to the proposal must be approved by the Human Research Ethics Office before they are implemented (except where an amendment is undertaken to eliminate an immediate risk to participants)
4. An annual progress report must be submitted to the Human Research Ethics Office on or before the anniversary of approval and a completion report submitted on completion of the project
5. Personnel working on this project must be adequately qualified by education, training and experience for their role, or supervised
6. Personnel must disclose any actual or potential conflicts of interest, including any financial or other interest or affiliation, that bears on this project
7. Changes to personnel working on this project must be reported to the Human Research Ethics Office
8. Data and primary materials must be retained and stored in accordance with the Western Australian University Sector Disposal Authority (WAUSDA) and the Curtin University Research Data and Primary Materials policy
9. Where practicable, results of the research should be made available to the research participants in a timely and clear manner
10. Unless prohibited by contractual obligations, results of the research should be disseminated in a manner that will allow public scrutiny; the Human Research Ethics Office must be informed of any constraints on publication
11. Approval is dependent upon ongoing compliance of the research with the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research, applicable legal requirements, and with Curtin University policies, procedures and governance requirements
12. The Human Research Ethics Office may conduct audits on a portion of approved projects.

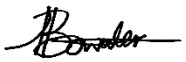
**Special Conditions of Approval**

Remove student mobile from recruitment email prior to sending.

**This letter constitutes low risk/negligible risk approval only**. This project may not proceed until you have met all of the Curtin University research governance requirements.

Should you have any queries regarding consideration of your project, please contact the Ethics Support Officer for your faculty or the Ethics Office at hrec@curtin.edu.au or on 9266 2784.

Yours sincerely



Amy Bowater
Acting Manager, Research Integrity

# Participant Consent Form




**CONSENT FORM**

| | |
|---|---|
| **HREC Project Number:** | HRE2017-0695 |
| **Project Title:** | An exploration of the association between movement and pain-related thoughts, beliefs and emotions in knee osteoarthritis. |
| **Chief Investigator:** | Dr Amity Campbell |
| **Student researcher:** | Tara West (PhD Candidate) |
| **Contact details:** | Tel: 9266 1001 or email: Tara.West@postgrad.curtin.edu.au |
| **Version:** | Version 2: 13/02/2018 |

- I have read the Participant Information Statement (version 2) and I understand its contents.
- I understand the purpose, extent and possible risks of my involvement in this project.
- I voluntarily consent to take part in this research project.
- I have had an opportunity to ask questions and I am satisfied with the answers I have received.
- I understand that my involvement is voluntary and I can withdraw at any time without problem.
- I understand that this project has been approved by Curtin University Human Research Ethics Committee and will be carried out in line with the National Statement on Ethical Conduct in Human Research (2007, updated March 2014).
- I understand I will receive a copy of this Information Statement and the Consent Form.

| | |
|---|---|
| Participant Name | |
| Participant Signature | |
| Date | |

Declaration by researcher: I have supplied an Information Letter and Consent Form to the participant who has signed above, and believe that they understand the purpose, extent and possible risks of their involvement in this project.

| | |
|---|---|
| Researcher Name | |
| Researcher Signature | |
| Date | |

***Note: All parties signing the Consent Form must date their own signature.***


Curtin University is a trademark of Curtin University. CRICOS Provider Code 00301J

# Participant Information Sheet

Curtin University

Thoughts, beliefs, emotions and movement in knee OA

**PARTICIPANT INFORMATION STATEMENT**

| | |
|---|---|
| **HREC Project Number:** | HRE2017-0695 |
| **Project Title:** | An exploration of the association between movement and pain-related thoughts, beliefs and emotions in knee osteoarthritis. |
| **Chief Investigator:** | Dr Amity Campbell, *PhD Supervisor* |
| **Co-Investigators:** | Dr Leo Ng, *PhD Co-supervisor*<br>Associate Professor Anne Smith, *PhD Associate Supervisor*<br>Professor Peter O'Sullivan, *PhD Associate Supervisor*<br>Associate Professor Peter Kent, *PhD Associate Supervisor* |
| **Student researcher:** | Tara Binnie, *PhD Candidate* |
| **Contact details** | Tel: 9266 1001 or email: Tara.Binnie@postgrad.curtin.edu.au |
| **Version number** | Version 4 |
| **Version date:** | 15/01/2019 |

**What is the Project About?**

There is growing evidence that peoples' thoughts, beliefs and emotions influence their movement - more specifically, their quality of movement *(the manner in which they perform a particular task)* and their movement quantity *(how much they perform or avoid a particular task)*. A better understanding of this link in people with knee OA will assist in the development of treatments that jointly target thoughts, beliefs, emotions and movement, which in turn may result in greater reductions in long-term pain and disability.

This study has three main aims:

- To investigate the usefulness of wearable movement sensors in measuring lower limb movement in individuals with knee osteoarthritis, during a number of physical tasks;
- To explore the links between pain-related thoughts, beliefs and emotions (via scores on questionnaires) and movement in individuals with knee osteoarthritis; and
- To develop a better understanding of the thoughts, beliefs and emotions held by people with knee osteoarthritis.

We are looking for 50 volunteers with knee osteoarthritis to participate in this study.


Curtin University is a trademark of Curtin University. CRICOS Provider Code 00301J

**Who is doing the Research**?
The study is being conducted by Tara Binnie at the School of Physiotherapy and Exercise Science at Curtin University. The results of this research project will be part of the work used by Tara Binnie to obtain a Doctor of Philosophy at Curtin University. This project is being funded by the University. There will be no costs to you and you will not be paid for participating in this project.

**Why am I being asked to take part and what will I have to do?**
You have been asked to take part because you have been diagnosed with knee osteoarthritis.
This a summary of what you will have to do, if you agree to participate:

At Curtin University (Kent Street, Bentley 6102):

- Attend a single data collection session at the Curtin University Motion Analysis Laboratory (maximum **2 hours**).
- Complete a number of questionnaires assessing pain-related thoughts, beliefs and emotions (approximately 15 minutes).
- Participate in a 20 minute interview to further explore your thoughts, beliefs and emotions surrounding your knee osteoarthritis. If you consent, the interview will be audio- and video-recorded so that the interviewer can concentrate on what you have to say and not be distracted by taking notes. After the interview, the interviewer will make a full written copy of the recording.
- Complete a number of physical tasks within this session, including walking a short distance, getting out of a chair, standing on one leg, step up, step down and walking up and down three stairs. You can use your normal walking aid and a rail as needed during these tasks. These tasks will be video-recorded.
- Have your movement measured within this session by wearing small wireless movement sensors on your thighs and lower legs.

At home:

- Continue to wear those small wireless movement sensors on your thighs and lower legs for the remainder of the day, after attending the data collection session at Curtin University.
- Have your movement quantity (activity) measured by wearing a small wireless activity monitor on your thigh for one week after attending the data session at Curtin University.

We will provide you with two reply postage paid envelopes, and will ask you to post the movement sensors and the activity monitor back to Curtin University after the required collection period. There will be no cost to you for taking part in this research. We will cover your car parking costs while you attend the session at Curtin University and will provide you with a $20 Coles Myer gift voucher to thank you for participating in our research.

**Are there any benefits' to being in the research project?**
The short interview will provide you with the opportunity to express your opinions and feelings about your knee and your diagnosis, and it is our experience that participants appreciate the opportunity to be listened to and provide their perspective of their condition. Furthermore, this project aims to explore how thoughts, beliefs and emotions influence an individual's behavioural responses to pain (specifically their movement quality and quantity). Understanding this link in people with knee OA will assist the development of treatments to jointly target beliefs, emotions and movement, which in turn may result in better long-term reductions in pain and disability. You may benefit from the findings of this research in the future.


Curtin University is a trademark of Curtin University. CRICOS Provider Code 00301J

**Are there any risks, side-effects or inconveniences from being in the research project?**

The risks associated with participating in this project are considered to be low. You might experience some physical discomfort or an exacerbation of your symptoms during the tasks or after the session, but the likelihood is low. That is because the tasks that will be undertaken are consistent with those routinely performed as part of a physiotherapy assessment and are commonly performed activities of daily living. Therefore, these measures are unlikely to put you at risk of harm. Furthermore, the session will be conducted by the primary researcher Tara Binnie, who is a registered physiotherapist and has experience working with individuals with knee osteoarthritis. The wireless movement sensors will be attached to your skin with hypo-allergenic adhesive tape, and although unlikely, this may cause minor skin irritation in some people. You will be required to complete a number of questionnaires and participate in a short interview to explore your thoughts, beliefs and emotions surrounding your knee(s). This may lead to emotional distress in some participants, however it is our experience that participants appreciate the opportunity to be listened to and provide their perspective of their condition. Should any emotional distress occur, the researcher will provide assistance and refer you to your GP and/or a counsellor.

**Who will have access to my information?**

The information collected in this research will be re-identifiable (coded). This means that we will remove personally identifying information (name, address etc.) on all data or samples and replace it with a code. Only the research team have access to the code to match your name if it is necessary to do so. Any information we collect will be treated as confidential and used only in this project unless otherwise specified. The following people will have access to the information we collect in this research: the research team and, in the event of an audit or investigation, staff from the Curtin University Office of Research and Development. Electronic data will be password-protected and hard copy data (including video or audio tapes) will be in locked storage. The information we collect in this study will be kept under secure conditions at Curtin University for 7 years after the research has ended and then it will be destroyed. The results of this research may be presented at conferences or published in professional journals. You will not be identified in any results that are published or presented.

**Will you tell me the results of the research?**

We will write to you at the end of the research (in about 18 months) and let you know the results of the research. Results will not be individual but based on all the information we collect and review as part of the research. The results of this research may be presented at conferences or published in professional journals. You will not be identified in any results that are published or presented. We will alert you of any publication of the results of the research by email.

**Do I have to take part in the research project?**

Taking part in a research project is voluntary. It is your choice to take part or not. You do not have to agree if you do not want to. If you decide to take part and then change your mind, that is okay, you can withdraw from the project. You do not have to give us a reason; just tell us that you want to stop. Please let us know you want to stop so we can make sure you are aware of any thing that needs to be done so you can withdraw safely. If you choose not to take part or start and then stop the study, it will not affect your relationship with the University, staff or colleagues. If you chose to leave the study we will use any information collected unless you tell us not to.


Curtin University is a trademark of Curtin University. CRICOS Provider Code 00301J

**What happens next and who can I contact about the research?**

If you decide to take part in this research we will ask you to sign the consent form. Signing it is telling us that you understand what you have read and what has been discussed. Signing the consent form indicates that you agree to be in the research project and have your health information used as described. Please take your time and ask any questions you have before you decide what to do. You will be given a copy of this information and the consent form to keep.

If you would like more information about the project, please contact the principal researcher Tara Binnie (PhD Candidate) on 9266 1001 or at Tara.Binnie@postgrad.curtin.edu.au

*Curtin University Human Research Ethics Committee (HREC) has approved this study (HRE2017-0695). Should you wish to discuss the study with someone not directly involved, in particular, any matters concerning the conduct of the study or your rights as a participant, or you wish to make a confidential complaint, you may contact the Ethics Officer on (08) 9266 9223 or the Manager, Research Integrity on (08) 9266 7093 or email hrec@curtin.edu.au.*


Curtin University is a trademark of Curtin University. CRICOS Provider Code 00301J

# Permissions to Adapt Published Material

Figure 2-2. Adapted with permission:
"A typical machine learning system" Adapted from " Machine Learning in Knee Osteoarthritis: A Review" by C. Kokkotis, S. Moustakidis, E. Papageorgiou, G. Giakas and D.E. Tsaopoulos, 2020, Osteoarthritis and Cartilage Open, 2(23), p. 2. (https://doi.org/10.1016/j.ocarto.2020.100069). Copyright 2020 by Osteoarthritis Research Society International (Creative Commons).

Dear Jay-Shian Tan,

Thank you for writing to us.

Kindly confirm if your usage is non-commercial as the article is under https://creativecommons.org/licenses/by-nc-nd/4.0/. If it is non-commercial permission is not required , kindly cite with correct citation.

Many thanks!

Kind regards,

**Kaveri Thakuria**
Senior Copyrights Coordinator
**ELSEVIER** | HCM - Health Content Management

Visit Elsevier Permissions

**From:** sculkin@ahint.com on behalf of Osteoarthritis Research Society International
**To:** Jay-Shian Tan
**Subject:** Re: Permissions for paper in Osteoarthritis and Cartilage Open
**Date:** Wednesday, 12 October 2022 12:24:08 AM

Hi Jay-Shian,

You have permission to use a modified version of a table and figure from a paper in Osteoarthritis and Cartilage Open for your thesis dissertation. Please let me know if you have any other questions. Thank you!

Best wishes,
Summer